



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 169374

TO: Bao-Qun Li

Location:

Art Unit: 1648

Thursday, November 03, 2005

Case Serial Number: 09/980559

From: Edward Hart

Location: Biotech-Chem Library

REM-1A55

Phone: 571-272-2512

edward.hart@uspto.gov

Search Notes

Examiner Li,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart



This Page Blank (uspto)

169374

STIC-Biotech/ChemLib

From: Li, Bao-Qun
Sent: Monday, October 24, 2005 1:51 PM
To: STIC-Biotech/ChemLib
Subject: 09/980,559

Please do the nucleic acid sequence homology and interference searches against amino acid sequence of SEQ ID NO: 2 and nucleic acid sequence of SEQ ID NO: 1, especially from nucleotide 341-9439 of SEQ ID NO: 1. Thanks.

Bao Qun Li M.D
TC 1600
Art Unit 1648
Tel. 517-272-0904
REM, 3C18
Rm. 3D24

Searcher: Doan
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: 11/2/05
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# 1 AA# 1
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

This Page Blank (uspto)



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact:*

Mary Hale, Information Branch Supervisor
Remsen Bldg. 01 D86
571-272-2507

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library Remsen Bldg.



This Page Blank (uspto)

Db 521 AGGCGCCAGCCCATCCCTTAAAGATCGCGCTCCACTGCGCAAACTCGGGGAAAAACAGGA 580
Oy 241 TACCCTTGCCCTATATACGGGAATGAGGGGACTCGGCTGCGAGGATGGCTCTGTCCGCC 300
Db 581 TACCCTTGCCCTATATACGGGAATGAGGGGACTCGGCTGCGAGGATGGCTCTGTCCGCC 640
Oy 301 CGAGGTTCCGCTCCCTCTTGGGAGCCCAATGACCCCGGCAATAGGTGCGCAACGTGGGT 360
Db 641 CGAGGTTCCGCTCCCTCTTGGGAGCCCAATGACCCCGGCAATAGGTGCGCAACGTGGGT 700
Oy 361 AAGGTCATCGATACCCTTAAAGTGGGCTTTGGCCGACCTGATGGGGTACAACCCCTGTGTG 420
Db 701 AAGGTCATCGATACCCTTAAAGTGGGCTTTGGCCGACCTGATGGGGTACAACCCCTGTGTG 760
Oy 421 GGGCGCCCGCTCGCGCGCGCTGCGCAGAGCTCTCGCGCATGGCGTGAAGTCTTGGAGAC 480
Db 761 GGGCGCCCGCTCGCGCGCGCTGCGCAGAGCTCTCGCGCATGGCGTGAAGTCTTGGAGAC 820
Oy 481 GGGGTTAATTTTGCACAGGGAATTACCCGAGTTGCTCTTTTCTATCTTCTGTGACC 540
Db 821 GGGGTTAATTTTGCACAGGGAATTACCCGAGTTGCTCTTTTCTATCTTCTGTGACC 880
Oy 541 CTGCTGCTGCTGATACCAACCCCGGCTCCGCTGCGCAAGTGAAGAACATAGTACCGGC 600
Db 881 CTGCTGCTGCTGATACCAACCCCGGCTCCGCTGCGCAAGTGAAGAACATAGTACCGGC 940
Oy 601 TACATGATGACTTAAAGACTGCAACCAATGACAGACTTACCTGGAGCTTCAAGCTGTGTC 660
Db 941 TACATGATGACTTAAAGACTGCAACCAATGACAGACTTACCTGGAGCTTCAAGCTGTGTC 1000
Oy 661 CTCCAGTCCCGGAGTGCCTCCGCTGCGAGAAAGTGGGAAATGCATCTCACTGTGATTA 720
Db 1001 CTCCAGTCCCGGAGTGCCTCCGCTGCGAGAAAGTGGGAAATGCATCTCACTGTGATTA 1060
Oy 721 CCGGCTGACCGGAATGTCGCTGCGAGGCGCCCGGCGCTCAGCAGGGCTTTCGAGACG 780
Db 1061 CCGGCTGACCGGAATGTCGCTGCGAGGCGCCCGGCGCTCAGCAGGGCTTTCGAGACG 1120
Oy 781 CACATCGACATGTTGTGATGTCGCGCAGCTCTGCTCTGCTCTTACGTCGAGGAGACTC 840
Db 1121 CACATCGACATGTTGTGATGTCGCGCAGCTCTGCTCTGCTCTTACGTCGAGGAGACTC 1180
Oy 841 TCGGCTGGGGTGAATGCTGCGACGCCAAATGTTTCAATGTTCTGCGCGACCACTGGTT 900
Db 1181 TCGGCTGGGGTGAATGCTGCGACGCCAAATGTTTCAATGTTCTGCGCGACCACTGGTT 1240
Oy 901 GTCCAGAGCTGCAATGCTGCACTCAACCTGGTACATCACTGACACCGGATGGCAATGG 960
Db 1241 GTCCAGAGCTGCAATGCTGCACTCAACCTGGTACATCACTGACACCGGATGGCAATGG 1300
Oy 961 GACATGATGATGAACCTGCTGCGCCACGAGCTACATGATCTTGGGCTACGCGATGCTGTC 1020
Db 1301 GACATGATGATGAACCTGCTGCGCCACGAGCTACATGATCTTGGGCTACGCGATGCTGTC 1360
Oy 1021 CCGGAGTCAATTAATGACATCAATGCGGGCTCATTTGGGCGCTGATGTTGGCTTGGCC 1080
Db 1361 CCGGAGTCAATTAATGACATCAATGCGGGCTCATTTGGGCGCTGATGTTGGCTTGGCC 1420
Oy 1081 TACTTCTCTATGACAGGAGCGTGGGCGAAAGTCGTTGTCATCTTGTGTTGGCGCGGG 1140
Db 1421 TACTTCTCTATGACAGGAGCGTGGGCGAAAGTCGTTGTCATCTTGTGTTGGCGCGGG 1480
Oy 1141 GTGAGCGCGCGACCCATCTGTTGGGGTTCGCGCGCAGACCAACCGGCGCTCAACC 1200
Db 1481 GTGAGCGCGCGACCCATCTGTTGGGGTTCGCGCGCAGACCAACCGGCGCTCAACC 1540
Oy 1201 AGCTTAATTTGACATGGGCCCCAGGCGAATAATCCAGCTCTGTTAACCAATGGCAGCTGG 1260
Db 1541 AGCTTAATTTGACATGGGCCCCAGGCGAATAATCCAGCTCTGTTAACCAATGGCAGCTGG 1600
Oy 1261 CACATCAACCGCAGCGCTGGAATGCAATGATCTCTTGGACACCGGCTTTATGCGCTCT 1320
Db 1601 CACATCAACCGCAGCGCTGGAATGCAATGATCTCTTGGACACCGGCTTTATGCGCTCT 1660

Oy 1321 CTGTTCTACACCCACAGCTTCAACTGTCGAGAGTGTCCCGAAAGCATGTCCGCTGCGCC 1380
Db 1661 CTGTTCTACACCCACAGCTTCAACTGTCGAGAGTGTCCCGAAAGCATGTCCGCTGCGCC 1720
Oy 1381 AGTATGAGAGGCTTCCGGGTGGGATGGGCGCTTGCATATAGAGATTAATGTACCAAT 1440
Db 1721 AGTATGAGAGGCTTCCGGGTGGGATGGGCGCTTGCATATAGAGATTAATGTACCAAT 1780
Oy 1441 CCAGAGATATAGAGCCCTATGCTGGCACTACCCAGAGGAGAGTGGGCGTGTCTCC 1500
Db 1781 CCAGAGATATAGAGCCCTATGCTGGCACTACCCAGAGGAGAGTGGGCGTGTGTCTCC 1840
Oy 1501 GCGAAGACTGTGTGGGCCAGTGTACTGTTTCAACCCCGACCCAGTGTGTGGGACAG 1560
Db 1841 GCGAAGACTGTGTGGGCCAGTGTACTGTTTCAACCCCGACCCAGTGTGTGGGACAG 1900
Oy 1561 ACCGACAGGCTTGGAGCGCCCACTTACAGTGGGGGAGAAATGAGCAATGTCTTCTTA 1620
Db 1901 ACCGACAGGCTTGGAGCGCCCACTTACAGTGGGGGAGAAATGAGCAATGTCTTCTTA 1960
Oy 1621 TTGAACAGCATGTCAGCACCGCTGGGGTCAATGAGTGGCTGACGTGATGAATCTTCT 1680
Db 1961 TTGAACAGCATGTCAGCACCGCTGGGGTCAATGAGTGGCTGACGTGATGAATCTTCT 2020
Oy 1681 GGCTACACCAAGACTTGGCGCGCACCAACCTGCGTACTAGAGCTGACTTCAACGCGACG 1740
Db 2021 GGCTACACCAAGACTTGGCGCGCACCAACCTGCGTACTAGAGCTGACTTCAACGCGACG 2080
Oy 1741 ACGGACTGTGTGGGCCACAGGACTGTTTAAAGAACATCTCTATACCACTTACCTTCAA 1800
Db 2081 ACGGACTGTGTGGGCCACAGGACTGTTTAAAGAACATCTCTATACCACTTACCTTCAA 2140
Oy 1801 TGGGCTCTGGGCGCTGGGCTCAGCGCAAGTGCCTGATGACTACCCCTACAGGCTCTGG 1860
Db 2141 TGGGCTCTGGGCGCTGGGCTCAGCGCAAGTGCCTGATGACTACCCCTACAGGCTCTGG 2200
Oy 1861 CATTACCCCTGACAGCTTAACTATACATCTTCAAAATAAGATGATGGAGGGGTT 1920
Db 2201 CATTACCCCTGACAGCTTAACTATACATCTTCAAAATAAGATGATGGAGGGGTT 2260
Oy 1921 GAGCAGAGGCTCAGCGCTGATGCAATTTCACTGTGGGAGATGGTTGCCAATGGAGAGC 1980
Db 2261 GAGCAGAGGCTCAGCGCTGATGCAATTTCACTGTGGGAGATGGTTGCCAATGGAGAGC 2320
Oy 1981 AGAGACAGAACTCAACTGTCTCTTGTGTCACCTGACCAACGGAATGGGCAATTTAACT 2040
Db 2321 AGAGACAGAACTCAACTGTCTCTTGTGTCACCTGACCAACGGAATGGGCAATTTAACT 2380
Oy 2041 TGCTCTTACTCGGACCTGCGCGCTTGTGATGTTCTTCTGCACTCCACCAAAACATC 2100
Db 2381 TGCTCTTACTCGGACCTGCGCGCTTGTGATGTTCTTCTGCACTCCACCAAAACATC 2440
Oy 2101 GTGAGGTAATTAATTCATGATGAGCTATACCTGCGCTCACAATAATACATCGTCGATG 2160
Db 2441 GTGAGGTAATTAATTCATGATGAGCTATACCTGCGCTCACAATAATACATCGTCGATG 2500
Oy 2161 GAGTGGGTAACTCTTATCTGCTCTTATGAGGGAAGCGGAGTTGGGCTGTATGG 2220
Db 2501 GAGTGGGTAACTCTTATCTGCTCTTATGAGGGAAGCGGAGTTGGGCTGTATGG 2560
Oy 2221 ATGCTCATCTTGTGGGCCAGGCGGAAGACAGCACTAAGAAAGCTGGTCACTTGGACGCT 2280
Db 2561 ATGCTCATCTTGTGGGCCAGGCGGAAGACAGCACTAAGAAAGCTGGTCACTTGGACGCT 2620
Oy 2281 GCGAGCGCAGCTAGCTGCAATGGCTCTATATTTTGTGATCTTTTTCGTCGTCGCTGG 2340
Db 2621 GCGAGCGCAGCTAGCTGCAATGGCTCTATATTTTGTGATCTTTTTCGTCGTCGCTGG 2680
Oy 2341 TACATCAAGGATCGGGATGCTCCCTTAGCTACCTATCTCCCTAGCTGGCTGTGGTCTTT 2400
Db 2681 TACATCAAGGATCGGGATGCTCCCTTAGCTACCTATCTCCCTAGCTGGCTGTGGTCTTT 2740

QY	2401	GGCCGACGCTCCTAGCAATGCCCAAGGCTTAATGCTTAAAGACGATCTGTGCATAGG	2460
Db	2741	AGCCGACTGCTCTAGCAATGCCCAAGGCTTAATGCTTAAAGACGATCTGTGCATAGG	2800
QY	2461	CAGATAGGAGGGGCTCTGCTGGTAATGATCACTCTTTACTCTCAACCCCGGGATTAAG	2520
Db	2801	CAGATAGGAGGGGCTCTGCTGGTAATGATCACTCTTTACTCTCAACCCCGGGATTAAG	2860
QY	2521	ACCCTTCTCAGCCGGTTTTTGTGGTGTGTGTGCTATCTTCTGACCCTGGGGAAGCTATG	2580
Db	2861	ACCCTTCTCAGCCGGTTTTTGTGGTGTGTGTGCTATCTTCTGACCCTGGGGAAGCTATG	2920
QY	2581	GTCCAGGAGTGGGGAACAATAAGAGGTGGCGGCTGGCCGTGAATGGAATGATAGGCC	2640
Db	2921	GTCCAGGAGTGGGGAACAATAAGAGGTGGCGGCTGGCCGTGAATGGAATGATAGGCC	2980
QY	2641	GTCCGCATATTTCACTCCAGAGTGGTGTGTTTGAATTAACAAGTGGCTTTTGGCGGTCTT	2700
Db	2981	GTCCGCATATTTCACTCCAGAGTGGTGTGTTTGAATTAACAAGTGGCTTTTGGCGGTCTT	3040
QY	2701	GGGCGCTGCTTACTTCTTAAAGTGTCTTGAACGCGGTGGCGTACTTGTCAAGGCTCAC	2760
Db	3041	GGGCGCTGCTTACTTCTTAAAGTGTCTTGAACGCGGTGGCGTACTTGTCAAGGCTCAC	3100
QY	2761	GCTCTACTGAGAGATGTGCACCATGCGCAAGGCATCTCGGGGGGCGAGTACGTCCAATG	2820
Db	3101	GCTCTACTGAGAGATGTGCACCATGCGCAAGGCATCTCGGGGGGCGAGTACGTCCAATG	3160
QY	2821	GGCGCTACTAGCCCTTGGCAGGTTGAGCTGGCACTTAACATCTAATGACACTCACCCCTATG	2880
Db	3161	GGCGCTACTAGCCCTTGGCAGGTTGAGCTGGCACTTAACATCTAATGACACTCACCCCTATG	3220
QY	2881	TCGGATTGGGCTGTAGTGTGGCTGGGACCTGGGCGGTGGCGGCTTGAAGCCTTATCATCTTC	2940
Db	3221	TCGGATTGGGCTGTAGTGTGGCTGGGACCTGGGCGGTGGCGGCTTGAAGCCTTATCATCTTC	3280
QY	2941	AGTCCGATGGAAGAAAGAAATGATTGTCTGGGGAAGCGAGACAAGCTGCTTGTGGGGAATTT	3000
Db	3281	AGTCCGATGGAAGAAAGAAATGATTGTCTGGGGAAGCGAGACAAGCTGCTTGTGGGGAATTT	3340
QY	3001	TTACACGGGACTTCCCGGTGCGGCCGAACTTGGTGGGAGGTCCTCTTGGCCGACGCTAT	3060
Db	3341	TTACACGGGACTTCCCGGTGCGGCCGAACTTGGTGGGAGGTCCTCTTGGCCGACGCTAT	3400
QY	3061	GGCTATATCTTCCAAAGGGGTGAGTCTTTCGCCCCCAATCACTGTGTTACGCCCAAGACA	3120
Db	3401	GGCTATATCTTCCAAAGGGGTGAGTCTTTCGCCCCCAATCACTGTGTTACGCCCAAGACA	3460
QY	3121	CGTGGCCTTTTGGGCAACATAGTGGTAGCATGACGGGGCGGACAGACAGAAACAGGCT	3180
Db	3461	CGTGGCCTTTTGGGCAACATAGTGGTAGCATGACGGGGCGGACAGACAGAAACAGGCT	3520
QY	3181	GGGGAATTTGAGTGTCTGTCCACAGTCACTAAGTCTTTCCTCGGAAATCATCATCTCGGG	3240
Db	3521	GGGGAATTTGAGTGTCTGTCCACAGTCACTAAGTCTTTCCTCGGAAATCATCATCTCGGG	3580
QY	3241	GTTTTGTGGAATGTGTCTACATGAGAGCTGGCAACAAGACTCTGGCCGGCTCAAGGGGTCCG	3300
Db	3581	GTTTTGTGGAATGTGTCTACATGAGAGCTGGCAACAAGACTCTGGCCGGCTCAAGGGGTCCG	3640
QY	3301	GTCAACGAGATGTACTCCAGTGTCTGAGGGGGACTTTAGTGGGTGGCCCAAGCCCCCTGGG	3360
Db	3641	GTCAACGAGATGTACTCCAGTGTCTGAGGGGGACTTTAGTGGGTGGCCCAAGCCCCCTGGG	3700
QY	3361	ACTTAATCTTTTGGAGCGGTGACAGTGTGGAACGGGTCCACTGTATCTGTGTACACGCGAAC	3420
Db	3701	ACTTAATCTTTTGGAGCGGTGACAGTGTGGAACGGGTCCACTGTATCTGTGTACACGCGAAC	3760
QY	3421	GCTGATGTCACTCCGGCTCGAAGACGCGGGGACAAACGGGGAGGCTTACTCTCCCGAGA	3480
Db	3761	GCTGATGTCACTCCGGCTCGAAGACGCGGGGACAAACGGGGAGGCTTACTCTCCCGAGA	3820
QY	3481	CCTCTTTTCCACTTTGAAGGGGTCTCAAGAGACCCCGGTGTATAGCCCCAGGGGCAAGCT	3540

Db	3821	CTCTTTTCACCTTGAAGGGGTCTCTCAGAGAGCCCGGTGCTATGCCACAGGGGCGACGCT	3880
Qy	3541	GTCCGAGCTCTCCGGGAGCGCTGTGTGCTCTCGGGGCGGTAGCTCATAGATTTTCATC	3600
Db	3881	GTCCGAGCTCTCCGGGAGCGCTGTGTGCTCTCGGGGCGGTAGCTCATAGATTTTCATC	3940
Qy	3601	CCGGTTAGACACTCGACATCTGTCACGCGGTCCCCACCTTTAGTGAACACAGACACCA	3660
Db	3941	CCCGTTAGACACTCGACATCTGTCAACCGGTCCCCACCTTTAGTGAACACAGACACCA	4000
Qy	3661	CCGTCTGTGCCCGACACTATCAGGTGCGGGTACTGTGATGCCCGGACCTGGCAGTGGGAAG	3720
Db	4001	CTGTCTGTGCCCGACACTATCAGGTGCGGGTACTGTGATGCCCGGACCTGGCAGTGGGAAG	4060
Qy	3721	AGCACCAAAAGTTCCTGTGCGATATGCTGCTCAGGGGTATAAAGTCTAGTGTCTTAATCCC	3780
Db	4061	AGCACCAAAAGTTCCTGTGCGATATGCTGCTCAGGGGTATAAAGTCTAGTGTCTTAATCCC	4120
Qy	3781	TCAGTGGCTGCGACCTCTGGGGTTTGGGCGCTACTGTTCTAAGGACATGGCATTAATCCC	3840
Db	4121	TCAGTGGCTGCGACCTCTGGGGTTTGGGCGCTACTGTTCTAAGGACATGGCATTAATCCC	4180
Qy	3841	AACAATTAGACATGGAATCAGAGACTGTGACGACCGGGGCGCCATCAGTACTCCACATAT	3900
Db	4181	AACAATTAGACATGGAATCAGAGACTGTGACGACCGGGGCGCCATCAGTACTCCACATAT	4240
Qy	3901	GGCAAAATTTCTCGCCGATGGGGGCTGTGCGGGCGGCGCTACAGACATCATCATATGTGAT	3960
Db	4241	GGCAAAATTTCTCGCCGATGGGGGCTGTGCGGGCGGCGCTACAGACATCATCATATGTGAT	4300
Qy	3961	GAATGCCATGCGGTGGACTCTTACCCATCTTGGCATCGGAACATGTCCTTGTATCAAGCA	4020
Db	4301	GAATGCCATGCGGTGGACTCTTACCCATCTTGGCATCGGAACATGTCCTTGTATCAAGCA	4360
Qy	4021	GAAGCAGCTGGGGTCAAGACTTAACTGTGTGGCTTACACTTACGCCCCCTGGGTCACTACA	4080
Db	4361	GAAGCAGCTGGGGTCAAGACTTAACTGTGTGGCTTACACTTACGCCCCCTGGGTCACTACA	4420
Qy	4081	ACCCCCACCCCAACATPAGAGAGAGTGGCCCTTGGGCGAGGAGGGCGAGATCCCTTCTAT	4140
Db	4421	ACCCCCACCCCAACATPAGAGAGAGTGGCCCTTGGGCGAGGAGGGCGAGATCCCTTCTAT	4480
Qy	4141	GGGAGGGCGAATTCCTCTGTCTTACATCAAGAGGAAAGAACATCTGATCTTCTGECATTC	4200
Db	4481	GGGAGGGCGAATTCCTCTGTCTTACATCAAGAGGAAAGAACATCTGATCTTCTGECATTC	4540
Qy	4201	AAAGAAAAAGTGTGACGAGCTTCGGGCGGGCCCTTGGGGGTATGGGCTTGAACCTCAGTGGCA	4260
Db	4541	AAAGAAAAAGTGTGACGAGCTTCGGGCGGGCCCTTGGGGGTATGGGCTTGAACCTCAGTGGCA	4600
Qy	4261	TACTACAGAGGGTTGAGCGCTCCCGTAAATACAACTCAGGGAAGACGTATGCTGTGCGCC	4320
Db	4601	TACTACAGAGGGTTGAGCGCTCCCGTAAATACAACTCAGGGAAGACGTATGCTGTGCGCC	4660
Qy	4321	ACCGACGCGCTCATGACAGGGTATACCTGGGAGCTTTGACTCCGTGATCGACTGCAACGTA	4380
Db	4661	ACCGACGCGCTCATGACAGGGTATACCTGGGAGCTTTGACTCCGTGATCGACTGCAACGTA	4720
Qy	4381	GGCGTCACTCAAGTTGTAGACTTCAAGTTTGAACCCCACTTACCATTAACACACAGATT	4440
Db	4721	GGCGTCACTCAAGTTGTAGACTTCAAGTTTGAACCCCACTTACCATTAACACACAGATT	4780
Qy	4441	GTCCCTCAAGACGTGTCTCAGCTAGGCAAGCGCGCGGGGTGCGACGGGTAGGGGAAGCTG	4500
Db	4781	GTCCCTCAAGACGTGTGTCTTCAAGTAAAGCAAGCGCGCGGGGTGCGACGGGTAGGGGAAGCTG	4840
Qy	4501	GGCAATTATAGGATATGTTTCCACTGTGTGACGAGCCTCAGAAATGTTTGAACATGTATGTG	4560
Db	4841	GGCAATTATAGGATATGTTTCCACTGTGTGTTCACATGGGTGACGAGCCTCAGAAATGTTTGAACATGTATGTG	4900
Qy	4561	CTCTGTAGTGTCTACGACGAGGGGCGCATGGTATGAGCTCACACCATGCGAAGCACACC	4620

D 4901 CTCTGTAGTGTCTACGACGAGGGGCGCATGTGTAGCTCACACCATCGAGAACACC 4960
Q 4621 GTCAAGCTCAGAGGCTATTTTGAACACGCGCGTTTGCTGTGTGCCAAGACCATCTTAG 4680
D 4961 GTCAAGCTCAGAGGCTATTTTGAACACGCGCGTTTGCTGTGTGCCAAGACCATCTTAG 5020
Q 4681 TTTTGGAGGAGGATTTTTCACCGGCTTCAACACATAGATGCTTCTTTTCCCAACA 4740
D 5021 TTTTGGAGGAGGATTTTTCACCGGCTTCAACACATAGATGCTTCTTTTCCCAACA 5080
Q 4741 AAGCAATGGGGGAAAAATTTGCACTTACAGCTTACGAGCTTACAGTGTGGCTAGG 4800
D 5081 AAGCAATGGGGGAAAAATTTGCACTTACAGCTTACAGCTTACAGTGTGGCTAGG 5140
Q 4801 GCCAAAGCCCCCCCCCGTCTGTGAGCGTCAATGTGAAGTGTGTGACTGCACTCAAGCCC 4860
D 5141 GCCAAAGCCCCCCCCCGTCTGTGAGCGTCAATGTGAAGTGTGTGACTGCACTCAAGCCC 5200
Q 4861 ACATCTGTGGGCCCCCAACCTTCTGTGACCGCTTGGGCTTGTTCACACAGAGTCAAC 4920
D 5201 ACATCTGTGGGCCCCCAACCTTCTGTGACCGCTTGGGCTTGTTCACACAGAGTCAAC 5260
Q 4921 CTCACACATCCCGTGAAGAAATPACATGCGCACTGCAATGCAAGCGCACTTGAAGTCAATG 4980
D 5261 CTCACACATCCCGTGAAGAAATPACATGCGCACTGCAATGCAAGCGCACTTGAAGTCAATG 5320
Q 4981 ACCAGACATGGGTCTTGGCAAGGGAGTCTTGGCGGCGCTGCGCGGTATTTGCTGCG 5040
D 5321 ACCAGACATGGGTCTTGGCAAGGGAGTCTTGGCGGCGCTGCGCGGTATTTGCTGCG 5380
Q 5041 ACCGGGTGTGTTCATATATGCGCGCTTGGCAATTTAACAGCGAGCGGTGTGGCGCG 5100
D 5381 ACCGGGTGTGTTCATATATGCGCGCTTGGCAATTTAACAGCGAGCGGTGTGGCGCG 5440
Q 5101 GACAAGAGGTCTCTATGAGGCTTTTATGAGATGAGGATGTGCTTACAGGCGGCT 5160
D 5441 GACAAGAGGTCTCTATGAGGCTTTTATGAGATGAGGATGTGCTTACAGGCGGCT 5500
Q 5161 CTATTTGAAGAGGGGACAGCGGATAGCCGAGATGCTGAAGTCCAAATCCAGGCTTATG 5220
D 5501 CTATTTGAAGAGGGGACAGCGGATAGCCGAGATGCTGAAGTCCAAATCCAGGCTTATG 5560
Q 5221 CAGCAAGCTTCCAAACAGGCTCAAGACATACACCACTGTGCAAGGCTTCAATGCGCAAG 5280
D 5561 CAGCAAGCTTCCAAACAGGCTCAAGACATACACCACTGTGCAAGGCTTCAATGCGCAAG 5620
Q 5281 GTAGAACAATTTCTGGGCGCAACACATGTGAACTTCAATTAGCGGATCCAAATCACTGCA 5340
D 5621 GTAGAACAATTTCTGGGCGCAACACATGTGAACTTCAATTAGCGGATCCAAATCACTGCA 5680
Q 5341 GGAATATCAACACTGCGCAAGGAAACCTTGCAAGTTCATGATGAGCGTTCAATGCGGCC 5400
D 5681 GGAATATCAACACTGCGCAAGGAAACCTTGCAAGTTCATGATGAGCGTTCAATGCGGCC 5740
Q 5401 CTCAACCAATCCGCTGTCAACAGCACTATCTTCTCAACATTTTGGGGGGCTGTGCTA 5460
D 5741 CTCAACCAATCCGCTGTCAACAGCACTATCTTCTCAACATTTTGGGGGGCTGTGCTA 5800
Q 5461 GCATCCCAATTTGACCAACCGCGGGGGCACTGGCTTGTGTCAAGTGTGCTGTAGGGA 5520
D 5801 GCATCCCAATTTGACCAACCGCGGGGGCACTGGCTTGTGTCAAGTGTGCTGTAGGGA 5860
Q 5521 GCTGCGTAGGCAATATAGGCTTAGATAGGTGCTAGTGAACATCTTGCGCAGGGATAGGT 5580
D 5861 GCTGCGTAGGCAATATAGGCTTAGATAGGTGCTAGTGAACATCTTGCGCAGGGATAGGT 5920
Q 5581 GCGGCGATTTTGGGGGGCTTGTGTGCAATTTCAAGATCATGTGTGGGAGAAACCTTCAATG 5640
D 5921 GCGGCGATTTTGGGGGGCTTGTGTGCAATTTCAAGATCATGTGTGGGAGAAACCTTCAATG 5980
Q 5641 GAGAGTGTGTCAACTTGTGCTGGAATTTCTGTCCGGGGGCTTGGTGTGTGGAGTCT 5700
D 5981 GAGAGTGTGTCAACTTGTGCTGGAATTTCTGTCCGGGGGCTTGGTGTGTGGAGTCT 6040

Q 5701 ATCTGCGGCGCAATTTCTGCGCCGACACGTGGAACCGGGGAAAGCGCGTCCAAATGATG 5760
D 6041 ATCTGCGGCGCAATTTCTGCGCCGACACGTGGAACCGGGGAAAGCGCGTCCAAATGATG 6100
Q 5761 AATGACTATTTGCTTTGCTTCCAGAGGAAATCAGTGGCCCCCACTCAATCTGTAGG 5820
D 6101 AATGACTATTTGCTTTGCTTCCAGAGGAAATCAGTGGCCCCCACTCAATCTGTAGG 6160
Q 5821 GAGTCGGAATGCGTGTGAGGCGTGTGACCCAACTACTTGGCTTCCCTTACATAACAGCTG 5880
D 6161 GAGTCGGAATGCGTGTGAGGCGTGTGACCCAACTACTTGGCTTCCCTTACATAACAGCTG 6220
Q 5881 CTCAGAAAGCTCCAACTGGAATTAATGAGGACTGCCCCATCCATGCGGCGCTGTGG 5940
D 6221 CTCAGAAAGCTCCAACTGGAATTAATGAGGACTGCCCCATCCATGCGGCGCTGTGG 6280
Q 5941 CTCGCGATGTGTGGGACTGTGGTTTGCACATCTCTTAAACAGCTTTAAAAATTTGGCTGACC 6000
D 6281 CTCGCGATGTGTGGGACTGTGGTTTGCACATCTCTTAAACAGCTTTAAAAATTTGGCTGACC 6340
Q 6001 TCCAAATTAATTCGCAAGATGCGCGGCTCCCTTTGTCTCTGTCAAAAAGGGTCAAG 6060
D 6341 TCCAAATTAATTCGCAAGATGCGCGGCTCCCTTTGTCTCTGTCAAAAAGGGTCAAG 6400
Q 6061 GCGGTGTGGCGCGCACTGGCATCATGACCAACGAGTGTCTTGGCGGCGCAATATCTCT 6120
D 6401 GCGGTGTGGCGCGCACTGGCATCATGACCAACGAGTGTCTTGGCGGCGCAATATCTCT 6460
Q 6121 GCGAATGTCCGCTTGGGCTTCATGAGAAATCAAGGGGCTTAAAGCTTGAATATCTGG 6180
D 6461 GCGAATGTCCGCTTGGGCTTCATGAGAAATCAAGGGGCTTAAAGCTTGAATATCTGG 6520
Q 6181 CAGGGACCTTTCCTATCAATTTGTTACAGGAGGCGCAGTGTGCGGAAACCGCGCGCA 6240
D 6521 CAGGGACCTTTCCTATCAATTTGTTACAGGAGGCGCAGTGTGCGGAAACCGCGCGCA 6580
Q 6241 AACTTTAAGTGTGCAATCTGAGAGGTGGCGGCTCAGAGTACCGGAGGTGACGACGAC 6300
D 6581 AACTTTAAGTGTGCAATCTGAGAGGTGGCGGCTCAGAGTACCGGAGGTGACGACGAC 6640
Q 6301 GGGTCATACCACTACATTAACAGGACTCACCACTGATTAACCTTGAATCCCTGCGCAACT 6360
D 6641 GGGTCATACCACTACATTAACAGGACTCACCACTGATTAACCTTGAATCCCTGCGCAACT 6700
Q 6361 CCTCTCCGAGTTCCTTCTGTGAGTGAACGAGTGCAGATCCATAGTTTGGCCCCACA 6420
D 6701 CCTCTCCGAGTTCCTTCTGTGAGTGAACGAGTGCAGATCCATAGTTTGGCCCCACA 6760
Q 6421 CCGAAGCGTTTTTCCGGGATGAGTCTGTGCTGCTGCGTTGGGCTTAATTCATTTGTCTCT 6480
D 6761 CCGAAGCGTTTTTCCGGGATGAGTCTGTGCTGCTGCGTTGGGCTTAATTCATTTGTCTCT 6820
Q 6481 GGGTCCGAGCTTCTTGGCGACCTTGAACCGGACACAGAGGTATTTAGTGTCAATGCTAAC 6540
D 6821 GGGTCCGAGCTTCTTGGCGACCTTGAACCGGACACAGAGGTATTTAGTGTCAATGCTAAC 6880
Q 6541 GATCATCTCATATCAACGCGGAGACTGAGCGCGGCTTTTGAAGCGGGGCTTCAACCCCA 6600
D 6881 GATCATCTCATATCAACGCGGAGACTGAGCGCGGCTTTTGAAGCGGGGCTTCAACCCCA 6940
Q 6601 TCCGAGGCAAGCTCTCGGAGACCACTATCGGCACATCGCTGCGAGCACACTGACCC 6660
D 6941 TCCGAGGCAAGCTCTCGGAGACCACTATCGGCACATCGCTGCGAGCACACTGACCC 7000
Q 6661 ACCCAAGGCAAAACCTATATATGAGGACTGTGACATGTGTGAAGTGTGCTTCAATGGGGGGGAT 6720
D 7001 ACCCAAGGCAAAACCTATATATGAGGACTGTGACATGTGTGAAGTGTGCTTCAATGGGGGGGAT 7060
Q 6721 GTGACTTCGATATAGTCTGGGCTCAAAAGTGTGCTTGTGACTCTCTCGACCCCAATGCTCT 6780
D 7061 GTGACTTCGATATAGTCTGGGCTCAAAAGTGTGCTTGTGACTCTCTCGACCCCAATGCTCT 7120

QY	6781	GAGAAAGAGCAGACTTGGACCTTGCATACATACAGAAATACATGCTCCCCAAGAGAG	6840
Db	7121	GAGAAAGAGCAGACTTGGACCTTGCATACATACAGAAATACATGCTCCCCAAGAGAG	7180
QY	6841	TTCCACCAAGCTTTACCGGCTTGGGCACAGGCTGTATTAACAACCAACGCTTGTGAAATCG	6900
Db	7181	TTCCACCAAGCTTTACCGGCTTGGGCACAGGCTGTATTAACAACCAACGCTTGTGAAATCG	7240
QY	6901	TGAAAAAGGCGAGATTACCAACCGGCGCATGTGCGGGCTGTGCTCCCTCCCTCAAGG	6960
Db	7241	TGAAAAAGGCGAGATTACCAACCGGCGCATGTGCGGGCTGTGCTCCCTCCCTCAAGG	7300
QY	6961	AAAACCCCGAGCGCTCCCCCAAGAGAGGCGCGGACAGTGGGCTTAAGTAGAGACTTCATTA	7020
Db	7301	AAAACCCCGAGCGCTCCCCCAAGAGAGGCGCGGACAGTGGGCTTAAGTAGAGACTTCATTA	7360
QY	7021	GGAATGCGCTTTCAACACAGCTGGCCATTAAAGTCTCTTTGGCCAGCCCCCCCCAGCGCGAT	7080
Db	7361	GGAATGCGCTTTCAACACAGCTGGCCATTAAAGTCTCTTTGGCCAGCCCCCCCCAGCGCGAT	7420
QY	7081	TCAGGCGCTTTCACAGGGGGGCGGGCGCTGCGCATTCGCGGCAGTGACAGCGCTCTGATGAG	7140
Db	7421	TCAGGCGCTTTCACAGGGGGGCGGGCGCTGCGCATTCGCGGCAGTGACAGCGCTCTGATGAG	7480
QY	7141	TTGGCCCTTTTCGAGACAGAGTTCCATCTCTTCCATGCCCCCCCCCTCGAGGGGAGACTTGA	7200
Db	7481	TTGGCCCTTTTCGAGACAGAGTTCCATCTCTTCCATGCCCCCCCCCTCGAGGGGAGACTTGA	7540
QY	7201	GATCCACACCTGGAGCGCTGAGCAGATGAGAGCCCCCAACCCCCCCCCAGGGGGGGTGGCA	7260
Db	7541	GATCCACACCTGGAGCGCTGAGCAGATGAGAGCCCCCAACCCCCCCCCAGGGGGGGTGGCA	7600
QY	7261	GCTCCCGGCTTCGAGCTCGGGGCTCTGTGTTACTTGTCTCGAGAGAGACGACTCGTCTGTG	7320
Db	7601	GCTCCCGGCTTCGAGCTCGGGGCTCTGTGTTACTTGTCTCGAGAGAGACGACTCGTCTGTG	7660
QY	7321	TGCTGCTTCATGTCAATACTCTTGGACCGGGGCTTTAATACTCTTGTAGTCCGGAAGAG	7380
Db	7661	TGCTGCTTCATGTCAATACTCTTGGACCGGGGCTTTAATACTCTTGTAGTCCGGAAGAG	7720
QY	7381	GAGAAAGTTACCGAATTAACCCCTTGAAGACATCCCTGTGAGATTCACAACAAGGTTAC	7440
Db	7721	GAGAAAGTTACCGAATTAACCCCTTGAAGACATCCCTGTGAGATTCACAACAAGGTTAC	7780
QY	7441	TGTACCAACAACAAGAGCGGCTCACTAAGGGCTTAAAGAGTAATCTTGTAGTAGATGCA	7500
Db	7781	TGTACCAACAACAAGAGCGGCTCACTAAGGGCTTAAAGAGTAATCTTGTAGTAGATGCA	7840
QY	7501	GTGCTGCATCTCTAATACGACTGATCTTTAAAGACATTAAGCTAAGGGGCTCCAGAGTC	7560
Db	7841	GTGCTGCATCTCTAATACGACTGATCTTTAAAGACATTAAGCTAAGGGGCTCCAGAGTC	7900
QY	7561	ACCGCAAGGCTCTCAACCATAGAGAGAGGCTTGCACAGTTAAACCCCAACCTTCTGCAGA	7620
Db	7901	ACCGCAAGGCTCTCAACCATAGAGAGAGGCTTGCACAGTTAAACCCCAACCTTCTGCAGA	7960
QY	7621	TCTTAATAATAGGTTTGGGGGCTAAGAGAGTCCGACGTTGTTCGGGAGGGGCGTTAAACAC	7680
Db	7961	TCTTAATAATAGGTTTGGGGGCTAAGAGAGTCCGACGTTGTTCGGGAGGGGCGTTAAACAC	8020
QY	7681	ATCAAGTCCGCTGAGAGAGACTCTCTGAGAGACTCAATAACCAATTTCCCAACAACATT	7740
Db	8021	ATCAAGTCCGCTGAGAGAGACTCTCTGAGAGACTCAATAACCAATTTCCCAACAACATT	8080
QY	7741	ATGGCCAAAAATGAGGTGTTCTGCTGTGACCCCAACCAAGGCGGCGAAGAAAGCAAGCTCGC	7800
Db	8081	ATGGCCAAAAATGAGGTGTTCTGCTGTGACCCCAACCAAGGCGGCGAAGAAAGCAAGCTCGC	8140
QY	7801	CTTAATCGTTTAAACCTGACCTCGAGGCTCGAGGCTTCGACAGAAAGATGGCCCTTTATGACATT	7860
Db	8141	CTTAATCGTTTAAACCTGACCTCGAGGCTCGAGGCTTCGACAGAAAGATGGCCCTTTATGACATT	8200
QY	7861	ACACAAAACTTCTCAAGCGGTGATAGGGGCTTCTTAATGATTCAGTAATTCCTCCCGCT	7920

Db	8201	ACACAAAACCTTCTCAGGCGGTGATGGGGCTTCTTAAGATTCCAGTATTTCCCCGCT	8260
Oy	7921	CAGCGGGTAGAGTTTCTCTTGAAAGCATGGCGGAAAGAAAGGACCCTATGGCTTTTCG	7980
Db	8261	CAGCGGGTAGAGTTTCTCTTGAAAGCATGGCGGAAAGAAAGGACCCTATGGCTTTTCG	8320
Oy	7981	TATGATACCCGATGCTTTGTAACCTCAACCGTCACCTGAGAGACATCAAGCATGAGAGATCC	8040
Db	8321	TATGATACCCGATGCTTTGTAACCTCAACCGTCACCTGAGAGACATCAAGCATGAGAGATCC	8380
Oy	8041	ATATATCGGGCTGCTCTTGGCCGAGAGGCCCACTGCGCATATCACTGCTAACTGAG	8100
Db	8381	ATATATCGGGCTGCTCTTGGCCGAGAGGCCCACTGCGCATATCACTGCTAACTGAG	8440
Oy	8101	AGACTTTACGTGGAGGGGCTTATGTTTCAACAGAAAGGCCCAACCTGCGGGTACAGGCT	8160
Db	8441	AGACTTTACGTGGAGGGGCTTATGTTTCAACAGAAAGGCCCAACCTGCGGGTACAGGCT	8500
Oy	8161	TGCGCGGACGAGCGGGTGTCTCAACACTAGCATGGGAAACCACTCATGCTACGTGAAA	8220
Db	8501	TGCGCGGACGAGCGGGTGTCTCAACACTAGCATGGGAAACCACTCATGCTACGTGAAA	8560
Oy	8221	GCTTTAGCGGCTTGTAAAGCTGCAAGGATATGCGGCCCAATGCTGTATGCGGAT	8280
Db	8561	GCTTTAGCGGCTTGTAAAGCTGCAAGGATATGCGGCCCAATGCTGTATGCGGAT	8620
Oy	8281	GACTTGATGTCATCTCAAGAAACCGAGGGGACCGAGAGAGACGAGCGGAAACCTGAGGCC	8340
Db	8621	GACTTGATGTCATCTCAAGAAACCGAGGGGACCGAGAGAGACGAGCGGAAACCTGAGGCC	8680
Oy	8341	TTCAAGAGGCTATGACCAAGTATTCTGCCCCCTCTGTGATACCCGCCAGACCGAGTAT	8400
Db	8681	TTCAAGAGGCTATGACCAAGTATTCTGCCCCCTCTGTGATACCCGCCAGACCGAGTAT	8740
Oy	8401	GATCTGAGCTGATTAACAATCTTGTCTCTCAAAATGTGTCTGTGGCGGTGGGCCCAAGGC	8460
Db	8741	GATCTGAGCTGATTAACAATCTTGTCTCTCAAAATGTGTCTGTGGCGGTGGGCCCAAGGC	8800
Oy	8461	CGCGGCAATATCTACCTGACCGAGAACCTTACACTCCATGCGCGCGCTGCTGGAA	8520
Db	8801	CGCGGCAATATCTACCTGACCGAGAACCTTACACTCCATGCGCGCGCTGCTGGAA	8860
Oy	8521	ACAGTTAGACACTCCCTGTCAATTCAATGCTGGAAACATATCAATACGCCCGGACC	8580
Db	8861	ACAGTTAGACACTCCCTGTCAATTCAATGCTGGAAACATATCAATACGCCCGGACC	8920
Oy	8581	ATATGGGCTCGATGCTCTGATGACACACTTCTTCATTTCTCATGCTCAAGACAG	8640
Db	8921	ATATGGGCTCGATGCTCTGATGACACACTTCTTCATTTCTCATGCTCAAGACAG	8980
Oy	8641	CTGACACGAAACTCACTTTGAGATGTAACGAGGCGGTACTCCGTGAGTCCCTTGAC	8700
Db	8981	CTGACACGAAACTCACTTTGAGATGTAACGAGGCGGTACTCCGTGAGTCCCTTGAC	9040
Oy	8701	CTCCGAGTAAATTTGAAAGGTTACATGGGCTTGAAGCTTTTCTTGACACATACACT	8760
Db	9041	CTCCGAGTAAATTTGAAAGGTTACATGGGCTTGAAGCTTTTCTTGACACATACACT	9100
Oy	8761	CCCCACGAACTGACACGCGGTGCTTCAACCCCTCAGAAAACTTGGGGCGCACCCCTCAGA	8820
Db	9101	CCCCACGAACTGACACGCGGTGCTTCAACCCCTCAGAAAACTTGGGGCGCACCCCTCAGA	9160
Oy	8821	GCGTGAAGAGCGCGGGCACGTGCAATCAGGGCTCCCTCATCTCCGTGGGGGGAAGCG	8880
Db	9161	GCGTGAAGAGCGCGGGCACGTGCAATCAGGGCTCCCTCATCTCCGTGGGGGGAAGCG	9220
Oy	8881	GCGGTTTGGGTCGATATCTCTTCAATTGGGGCGGTGAAGACCAAGCTCAAACTCATCTCA	8940
Db	9221	GCGGTTTGGGTCGATATCTCTTCAATTGGGGCGGTGAAGACCAAGCTCAAACTCATCTCA	9280
Oy	8941	TTGCGGGAAGCGCGCTCTCTGAGTTTATCAAGTGTTCACCGTGGCGCGCGGGGCGC	9000

Db 9281 TTGCCGGAAGCGCCTCTCCGATTTATTCAGCTGTTCACCGTCGGCGCCGCGGCGG 9340
Qy 9001 GACATTTATCAGACGCTGTCCGCTGCCGACCCCGCTTATGTCTTTGGCTTACTCTTA 9060
Db 9341 GACATTTATCAGACGCTGTCCGCTGCCGACCCCGCTTATGTCTTTGGCTTACTCTTA 9400
Qy 9061 CTTTTTGTAGGGGTAGGCTTTTCTTACTCTCCGCTCGG 9099
Db 9401 CTTTTTGTAGGGGTAGGCTTTTCTTACTCTCCGCTCGG 9439

RESULT 2
AX057317
LOCUS AX057317 9711 bp DNA linear PAT 17-JAN-2001
DEFINITION Sequence 3 from Patent WO0075352.
ACCESSION AX057317
VERSION AX057317.1 GI:12310062
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.

REFERENCE
1
AUTHORS Nam, J.H., Bukh, J., Emerson, S.U. and Purcell, R.H.
TITLE Hcv/bvdr chimeric genomes and uses thereof
JOURNAL Patent: WO 0075352-A 3 14-DEC-2000;
THE GOVERNMENT OF THE UNITED STATES OF AMERICA (US)
FEATURES
SOURCE location/Qualifiers
1..9711
/organism="Hepatitis C virus"
/mol_type="unassigned DNA"
/db_xref="taxon:11103"

ORIGIN
Query Match 100.0%; Score 9099; DB 6; Length 9711;
Best local Similarity 100.0%; Pred. No. 0;
Matches 9099; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGCAAAATCTTAACTCAAGAAAACCAAAAGAAAACCAAAACCGTCCGCAAA 60
Db 341 ATGAGCAAAATCTTAACTCAAGAAAACCAAAAGAAAACCAAAACCGTCCGCAAA 400
Qy 61 GACGTTAAGTTTCCGCGCGCGCGCCAGATCGTTGGCGGAGTATCTTTGCCCGCAAG 120
Db 401 GACGTTAAGTTTCCGCGCGCGCGCCAGATCGTTGGCGGAGTATCTTTGCCCGCAAG 460
Qy 121 GAGCCGAGGTTGGGTTGGCGCGCGCAAGAACTTGAGCGGTTCCAGCGCACTGGA 180
Db 461 GAGCCGAGGTTGGGTTGGCGCGCGCAAGAACTTGAGCGGTTCCAGCGCACTGGA 520
Qy 181 AGGCGCAGCCCATCTTAAAGATCGCGCTCACTGCAAAATCTTGGGAAAAACAGAA 240
Db 521 AGGCGCAGCCCATCTTAAAGATCGCGCTCACTGCAAAATCTTGGGAAAAACAGAA 580
Qy 241 TACCCCTGGCCCTTATACGGAATGAGGAACTCGGCTGGGCAAGATGCTCTCTCC 300
Db 581 TACCCCTGGCCCTTATACGGAATGAGGAACTCGGCTGGGCAAGATGCTCTCTCC 640
Qy 301 CGAGGTTCCGCTCTTTGGGGGCCCAATGACCCCGCGCAATGAGTGGCGAAAGTGGGT 360
Db 641 CGAGGTTCCGCTCTTTGGGGGCCCAATGACCCCGCGCAATGAGTGGCGAAAGTGGGT 700
Qy 361 AAGGTCATGATACCTTAACGTGCGGCTTTGGCGCACTCATGAGGTACATCCTGTGTG 420
Db 701 AAGGTCATGATACCTTAACGTGCGGCTTTGGCGCACTCATGAGGTACATCCTGTGTG 760
Qy 421 GCGCGCCGCTGCGCGCGGTGCGCAGAGCTCTCGCGCATGCGGTGAGAGTCTTGAAGAC 480
Db 761 GCGCGCCGCTGCGCGCGGTGCGCAGAGCTCTCGCGCATGCGGTGAGAGTCTTGAAGAC 820
Qy 481 GGGGTTATTTTGAACAGGAACTTACCGGTTGCTCTTTTATCTTCTGTGGCC 540
Db 821 GGGGTTATTTTGAACAGGAACTTACCGGTTGCTCTTTTATCTTCTGTGGCC 880

Qy 541 CTGCTGTCTTGATCAACCAACCCCGTCTTCGCTGCCGAAGTGAAGAACTACGTAACCGGC 600
Db 881 CTGCTGTCTTGATCAACCAACCCCGTCTTCGCTGCCGAAGTGAAGAACTACGTAACCGGC 940
Qy 601 TACATGTGATCTTAAGCATGTCACCAATGACAGATTAAGTGGAGCTCCAGGCTGTGTC 660
Db 941 TACATGTGATCTTAAGCATGTCACCAATGACAGATTAAGTGGAGCTCCAGGCTGTGTC 1000
Qy 661 CTCACAGTCCCGGAGTCTCCGTCGAGGAAAGTGGGAAATGCAATCTGAGTGGATA 720
Db 1001 CTCACAGTCCCGGAGTCTCCGTCGAGGAAAGTGGGAAATGCAATCTGAGTGGATA 1060
Qy 721 CCGGTCTACCGAATGTGGCCGTGACGCGGCCCGGCGCTTACCGAGGCTTTGCGGACG 780
Db 1061 CCGGTCTACCGAATGTGGCCGTGACGCGGCCCGGCGCTTACCGAGGCTTTGCGGACG 1120
Qy 781 CACATGCAATGATGTTGATGTCGCCGACAGCTCTGCTGCTCTCACTGGGAGACCTTC 840
Db 1121 CACATGCAATGATGTTGATGTCGCCGACAGCTCTGCTGCTCTCACTGGGAGACCTTC 1180
Qy 841 TGGCGTGGGATGATGCTCGACGCCCAATGTTCAATGTTCTCGCCGACGACCACTGTT 900
Db 1181 TGGCGTGGGATGATGCTCGACGCCCAATGTTCAATGTTCTCGCCGACGACCACTGTT 1240
Qy 901 GTCCAAAGCTGCAATGTTCTCAATCTTACCTGATCCATCATGACATGACACCGCATGG 960
Db 1241 GTCCAAAGCTGCAATGTTCTCAATCTTACCTGATCCATCATGACATGACACCGCATGG 1300
Qy 961 GACATGATGATGAATGTCGTGCCCAAGCTTACCAATGTTGGCGGATGCGTGTGTC 1020
Db 1301 GACATGATGATGAATGTCGTGCCCAAGCTTACCAATGTTGGCGGATGCGTGTGTC 1360
Qy 1021 CCCGAGTCAATTAAGACATCAATTAAGCGGAGCTCATTTGGGAGCTCATGTTCCGCTTGGCC 1080
Db 1361 CCCGAGTCAATTAAGACATCAATTAAGCGGAGCTCATTTGGGAGCTCATGTTCCGCTTGGCC 1420
Qy 1421 TACTTCTATGAGGAGAGCTGTGGCGAAGGTGTCATCTTCTGTGGCGCGCGG 1480
Db 1481 TACTTCTATGAGGAGAGCTGTGGCGAAGGTGTCATCTTCTGTGGCGCGCGG 1540
Qy 1441 GTGGAAGCGCGCACTTACTGTTGGGGTTTCGCGCGCAAGCAACCGGGGCGCTCAC 1200
Db 1481 GTGGAAGCGCGCACTTACTGTTGGGGTTTCGCGCGCAAGCAACCGGGGCGCTCAC 1540
Qy 1201 AGCTTATTTGACATGAGGAGCTGTGGCGAAGGTGTCATCTTCTGTGGCGCGCGG 1260
Db 1541 AGCTTATTTGACATGAGGAGCTGTGGCGAAGGTGTCATCTTCTGTGGCGCGCGG 1600
Qy 1261 CACATCAACCGCACCGCCCTGAACTGCAATGACTCTTGCACACCGGCTTTATCGCTCT 1320
Db 1601 CACATCAACCGCACCGCCCTGAACTGCAATGACTCTTGCACACCGGCTTTATCGCTCT 1660
Qy 1321 CTGTTTCAACCCAGCTTCAACTGCTGAGATGTCGCCGAAGCAATGTCGCTGCGCC 1380
Db 1661 CTGTTTCAACCCAGCTTCAACTGCTGAGATGTCGCCGAAGCAATGTCGCTGCGCC 1720
Qy 1381 AGTATGAGGCTTCCGCGGTGAGTGGGCGCTTGCATATGAGGATTAATGTCACAAT 1440
Db 1721 AGTATGAGGCTTCCGCGGTGAGTGGGCGCTTGCATATGAGGATTAATGTCACAAT 1780
Qy 1441 CAGAGATATGAGACCTTATTTGCTGCACTACCAACCAAGCATGTTGGCTGTCTCC 1500
Db 1781 CAGAGATATGAGACCTTATTTGCTGCACTACCAACCAAGCATGTTGGCTGTCTCTCC 1840
Qy 1501 GCGAAGATGCTGTGTCGCCAGTGTACTGTTTACCCCAAGGCGCAGTGGTGGGACG 1560
Db 1841 GCGAAGATGCTGTGTCGCCAGTGTACTGTTTACCCCAAGGCGCAGTGGTGGGACG 1900
Qy 1561 ACCGACAGGCTTGAAGCGCCCACTTACAGCTGGGGGGAAGATGAGACAGATGCTTCTTA 1620
Db 1901 ACCGACAGGCTTGAAGCGCCCACTTACAGCTGGGGGGAAGATGAGACAGATGCTTCTTA 1960

QY 1621 TTGAAGACGCTGACACCGCTGGGGTCAATGTTCCGGCTGCACGATGATGAATCTCTT 1680
DB 1961 TTGAAGACGCTGACACCGCTGGGGTCAATGTTCCGGCTGCACGATGATGAATCTCTT 2020
QY 1681 GGCCTACCAAGACTTGGGGGCAACACCTGCGCTAATAGAGTGAATTTCAAGCCAGC 1740
DB 2021 GGCCTACCAAGACTTGGGGGCAACACCTGCGCTAATAGAGTGAATTTCAAGCCAGC 2080
QY 1741 AGGACCTGTTGGCCCAAGGACTGTTTGAAGAGATCCTATATACCTTACCTTCAA 1800
DB 2081 AGGACCTGTTGGCCCAAGGACTGTTTGAAGAGATCCTATATACCTTACCTTCAA 2140
QY 1801 TGGGCTCTGGGGCTTGGCTCAAGCCAAAGGTGCTGATGACATACCTTCAAGGCTTGG 1860
DB 2141 TGGGCTCTGGGGCTTGGCTCAAGCCAAAGGTGCTGATGACATACCTTCAAGGCTTGG 2200
QY 1861 CATTACCCCTGCACAGTTAACTATACCATCTTCAAAATAGATGATGTTGGAGGGGTT 1920
DB 2201 CATTACCCCTGCACAGTTAACTATACCATCTTCAAAATAGATGATGTTGGAGGGGTT 2260
QY 1921 GAGCAGAGGCTCAAGGCTGATGCAATTTGACTGTTGGGAGTGGTTGCACTTGGAGAG 1980
DB 2261 GAGCAGAGGCTCAAGGCTGATGCAATTTGACTGTTGGGAGTGGTTGCACTTGGAGAG 2320
QY 1981 AGAGACAGAGTCAAGTCTCTCTTGGTGCATCCACAGGAATGGGCAATTTTACT 2040
DB 2321 AGAGACAGAGTCAAGTCTCTCTTGGTGCATCCACAGGAATGGGCAATTTTACT 2380
QY 2041 TGGCTTTACTGGACCTGCGCGCTTGTGCACTGGTCTTCTGACCTTCAACAAACATC 2100
DB 2381 TGGCTTTACTGGACCTGCGCGCTTGTGCACTGGTCTTCTGACCTTCAACAAACATC 2440
QY 2101 GTGAGAGTCAATTTCAATGTAATGGCTATACCTGCGCTCAAAATATACCTGCGAGT 2160
DB 2441 GTGAGAGTCAATTTCAATGTAATGGCTATACCTGCGCTCAAAATATACCTGCGAGT 2500
QY 2161 GAGTGGTAAATCTCTTATCTCTGCTCTTACGGAAGCCAGGGTGGCGCTCTTATG 2220
DB 2501 GAGTGGTAAATCTCTTATCTCTGCTCTTACGGAAGCCAGGGTGGCGCTCTTATG 2560
QY 2221 ATGCTCATCTTGTGGGCGACGCGGAAGACGACTAGAGAAGCTGCTATCTTGCAGCT 2280
DB 2561 ATGCTCATCTTGTGGGCGACGCGGAAGACGACTAGAGAAGCTGCTATCTTGCAGCT 2620
QY 2281 GCGAGGCGAGTACGCTGCAATGCTTCTATATTTTGTCACTTTTTCGTGGCTGTGG 2340
DB 2621 GCGAGGCGAGTACGCTGCAATGCTTCTATATTTTGTCACTTTTTCGTGGCTGTGG 2680
QY 2341 TACATCAAGGGTGGGGATGCCCTTAGCTACCTATCCCTCACTGGCGTGGTCCCTT 2400
DB 2681 TACATCAAGGGTGGGGATGCCCTTAGCTACCTATCCCTCACTGGCGTGGTCCCTT 2740
QY 2401 AGGCTTCTGCTAGCAATTTGCCCAACAGGCTTATGCTTATGACGATCTGTCATG 2460
DB 2741 AGGCTTCTGCTAGCAATTTGCCCAACAGGCTTATGCTTATGACGATCTGTCATG 2800
QY 2461 CAGATGAGAGGCTCTGCTGTAATGATCACTCTTATCTTCACTGCACTCCCGGATTAAG 2520
DB 2801 CAGATGAGAGGCTCTGCTGTAATGATCACTCTTATCTTCACTGCACTCCCGGATTAAG 2860
QY 2521 ACCCTTCTAGCCGGTTTTGTGGTGGTGTGCTATCTTGAACCTGAGGGGAAGCTATG 2580
DB 2861 ACCCTTCTAGCCGGTTTTGTGGTGGTGTGCTATCTTGAACCTGAGGGGAAGCTATG 2920
QY 2581 GTTCAGAGTGGGCACTTATGCAAGTGCAGGTGGCGGTATGATGATGATGATGATG 2640
DB 2921 GTTCAGAGTGGGCACTTATGCAAGTGCAGGTGGCGGTATGATGATGATGATGATG 2980
QY 2641 GTGCAATATTTCAACAGGTGGTGTGTTGATATACCAAGGGCTCTTGGGGGTGCTT 2700
DB 2981 GTGCAATATTTCAACAGGTGGTGGTGTGTTGATATACCAAGGGCTCTTGGGGGTGCTT 3040
QY 2701 GGGCTGTCTTACTCTTAAAGGTGGTGTGACGCGGTGCCGTACTTCCGTCAGGGCTCAC 2760

DB 3041 GGGCTGTCTTACTCTTAAAGGTGGTGTGACGCGGTGCCGTACTTCCGTCAGGGCTCAC 3100
QY 2761 GCTTCTAGAGATGTGACCATGCAAGGACTTCTGGGGGGGAGATGCTCAAGT 2820
DB 3101 GCTTCTAGAGATGTGACCATGCAAGGACTTCTGGGGGGGAGATGCTCAAGT 3160
QY 2821 GGGCTGTAGCCCTTGGGAGGTGGAGTGGACTTATATGATACCACTTACCCCTATG 2880
DB 3161 GGGCTGTAGCCCTTGGGAGGTGGAGTGGACTTATATGATACCACTTACCCCTATG 3220
QY 2881 TCGAATGGGCTGTAGGCTGCGGAGACTTGGCGGTGCGCTTGAAGCTTATCTTCT 2940
DB 3221 TCGAATGGGCTGTAGGCTGCGGAGACTTGGCGGTGCGCTTGAAGCTTATCTTCT 3280
QY 2941 AGTCCATGGAAGAAAGTCAATGTTCTGGGAGCGGAGACAGCTGCTTGGGAGATT 3000
DB 3281 AGTCCATGGAAGAAAGTCAATGTTCTGGGAGCGGAGACAGCTGCTTGGGAGATT 3340
QY 3001 TTACAGGACTTCCGGTCCGCGGACTTGGTGGGAGGTCTCTTGGCCGAGCTGAT 3060
DB 3341 TTACAGGACTTCCGGTCCGCGGACTTGGTGGGAGGTCTCTTGGCCGAGCTGAT 3400
QY 3061 GGCCTATACCTCAAGGGGATGAGTCTTCCGCCCATCATCTTACGCCGAGAGACA 3120
DB 3401 GGCCTATACCTCAAGGGGATGAGTCTTCCGCCCATCATCTTACGCCGAGAGACA 3460
QY 3121 CGTGGCCCTTTGGGCACTATGTTGAGATGACGAGGCGGACAAAGACAGGCT 3180
DB 3461 CGTGGCCCTTTGGGCACTATGTTGAGATGACGAGGCGGACAAAGACAGGCT 3520
QY 3181 GGGGAAATTCAGGTCCTGTGCAAGTCACTGCTCTTCTGGGAATCATCTGGGG 3240
DB 3521 GGGGAAATTCAGGTCCTGTGCAAGTCACTGCTCTTCTGGGAATCATCTGGGG 3580
QY 3241 GTTTTGTGACCTGTCTTACCATGAGCTGCAACAAAGCTTGGCCGAGTCCG 3300
DB 3581 GTTTTGTGACCTGTCTTACCATGAGCTGCAACAAAGCTTGGCCGAGTCCG 3640
QY 3301 GTCAAGCAATGTAATCTCAAGTGTGAGAGGAGGACTTATGAGGTGCGCCAGCCCTGG 3360
DB 3641 GTCAAGCAATGTAATCTCAAGTGTGAGAGGAGGACTTATGAGGTGCGCCAGCCCTGG 3700
QY 3361 ACTAAATCTTGGAGCGGTGCAAGTGTGAGAGCGGTGACCTGATCCGTGACGCGGAA 3420
DB 3701 ACTAAATCTTGGAGCGGTGCAAGTGTGAGAGCGGTGACCTGATCCGTGACGCGGAA 3760
QY 3421 GCTGATGTCATCCGGCTGGAAGACGCGGGGACAAAGCGGAGGCTTATCTTCCGAGA 3480
DB 3761 GCTGATGTCATCCGGCTGGAAGACGCGGGGACAAAGCGGAGGCTTATCTTCCGAGA 3820
QY 3481 CCTCTTTCACCTTGAAGGGGTCTCAGAGAGCGGTGCTATGCCCCAGGGGCAAGCT 3540
DB 3821 CCTCTTTCACCTTGAAGGGGTCTCAGAGAGCGGTGCTATGCCCCAGGGGCAAGCT 3880
QY 3541 GTGAGAGTCTTCCGGGAGCTGTGCTCTCAGAGAGCGGTGCTATGCCCCAGGGGCAAGCT 3600
DB 3881 GTGAGAGTCTTCCGGGAGCTGTGCTCTCAGAGAGCGGTGCTATGCCCCAGGGGCAAGCT 3940
QY 3601 CCGGTGAGACACTGACATCTGTCAGCGGTGCCCACTTTATGTAACAACAGACACA 3660
DB 3941 CCGGTGAGACACTGACATCTGTCAGCGGTGCCCACTTTATGTAACAACAGACACA 4000
QY 3661 CCGGTGAGACACTGACATCTGTCAGCGGTGCCCACTTTATGTAACAACAGACACA 3720
DB 4001 CCGGTGAGACACTGACATCTGTCAGCGGTGCCCACTTTATGTAACAACAGACACA 4060
QY 3721 AGCACCAGAGTCTCTGTCATATGCTGCTCAGGGGATTAAGAGTCTTATGCTTATCC 3780
DB 4061 AGCACCAGAGTCTCTGTCATATGCTGCTCAGGGGATTAAGAGTCTTATGCTTATCC 4120
QY 3781 TCAAGTGTGCAACCTTGGGGTGGGGGCTTATGCTTAAAGGACATGACATCAATCC 3840

Db 4121 TCAGTGGCTGCCACCTTGAGGGTGTGGGCGTACTGTGCTAAGGCACATGGCATCAATCCC 4180
Qy 3841 AACATTAGGACTGAGAGTCAAGGACTGTGACGACCGGGGGGCCCATCAGACTCCACATAT 3900
Db 4181 AACATTAGGACTGAGAGTCAAGGACTGTGACGACCGGGGGGCCCATCAGACTCCACATAT 4240
Qy 3901 GGCAAATTCCTCGCCGATGGGGGCTGTGCGGGCGGCCCTTACGACATCATATATGTAT 3960
Db 4241 GGCAAATTCCTCGCCGATGGGGGCTGTGCGGGCGGCCCTTACGACATCATATATGTAT 4300
Qy 3961 GAATGCCATGGCCGTGAGACTCTTACCAACCATCTTGGGATCCGGAAGATCCCTTGAACGA 4020
Db 4301 GAATGCCATGGCCGTGAGACTCTTACCAACCATCTTGGGATCCGGAAGATCCCTTGAACGA 4360
Qy 4021 GAGACAGCTGGGGTCAAGACTAACTGTGCTGAGCTACAGCTACGCCCTTGAGTCAAGTACA 4080
Db 4361 GAGACAGCTGGGGTCAAGACTAACTGTGCTGAGCTACAGCTACGCCCTTGAGTCAAGTACA 4420
Qy 4081 ACCCCCCACCCCAACATAGAGAGGTGGCCCTTGGGCGAGAGGGCGAGATCCCTTCTAT 4140
Db 4421 ACCCCCCACCCCAACATAGAGAGGTGGCCCTTGGGCGAGAGGGCGAGATCCCTTCTAT 4480
Qy 4141 GGGAGGGGCGATTCCTCTGTCTTATCATCAAGGGAGGAAAGACATCTGATCTTCTGCGATTCA 4200
Db 4481 GGGAGGGGCGATTCCTCTGTCTTATCATCAAGGGAGGAAAGACATCTGATCTTCTGCGATTCA 4540
Qy 4201 AAGAAAAAGTGTGACGAGCTCGCGCGGCCCTTGGGGGTATGGGCTTGAATCTAGTGGCA 4260
Db 4541 AAGAAAAAGTGTGACGAGCTCGCGCGGCCCTTGGGGGTATGGGCTTGAATCTAGTGGCA 4600
Qy 4261 TACTACAGAGGGTGTGAGCGTCTCGGTAATACCACTCAGGGAGAGATAGTGTGTGCGCC 4320
Db 4601 TACTACAGAGGGTGTGAGCGTCTCGGTAATACCACTCAGGGAGAGATAGTGTGTGCGCC 4660
Qy 4321 ACCGACGCCCTTCAAGACGGGTATATCTGGGGACTTTGATCTCGGATGTGATGCAAGTGA 4380
Db 4661 ACCGACGCCCTTCAAGACGGGTATATCTGGGGACTTTGATCTCGGATGTGATGCAAGTGA 4720
Qy 4381 GCGGTCATCTCAAGTGTGTGACTTTCAGTTAGACCCACATTCACCATTAACCAACAGATT 4440
Db 4721 GCGGTCATCTCAAGTGTGTGACTTTCAGTTAGACCCACATTCACCATTAACCAACAGATT 4780
Qy 4441 GTCCCTCAAGACGCTGTCTCAAGTACGACGCGCGGGGTGCGACGGGTATGGGGAAAGTGTG 4500
Db 4781 GTCCCTCAAGACGCTGTCTCAAGTACGACGCGCGGGGTGCGACGGGTATGGGGAAAGTGTG 4840
Qy 4501 GGCATTTATAGGTATGTTCACACTGTGTGAGGAGGCTTACAGAAATTTTGAACAGTGTATG 4560
Db 4841 GGCATTTATAGGTATGTTCACACTGTGTGAGGAGGCTTACAGAAATTTTGAACAGTGTATG 4900
Qy 4561 CTCTGTAGTGTCTACGACGACGAGGGCGCATGTGTAGCTCACACCATCCGAAACACAC 4620
Db 4901 CTCTGTAGTGTCTACGACGACGAGGGCGCATGTGTAGCTCACACCATCCGAAACACAC 4960
Qy 4621 GTCAAGGCTCAAGGCGTATTTCAACACGCGCGGTTTCTGTGTGCAAGACCATCTTGTAG 4680
Db 4961 GTCAAGGCTCAAGGCGTATTTCAACACGCGCGGTTTCTGTGTGCAAGACCATCTTGTAG 5020
Qy 4681 TTTTGGGAGGCGATTTTCAACGCGGCTCACACATAGATGCCCATCTTCTTCCCAACA 4740
Db 5021 TTTTGGGAGGCGATTTTCAACGCGGCTCACACATAGATGCCCATCTTCTTCCCAACA 5080
Qy 4741 AAGCAATCGGGGGAAAAATTTGCGATCTTAACAGCTACAGGCTACAGGCTACAGTGTGCGCTAG 4800
Db 5081 AAGCAATCGGGGGAAAAATTTGCGATCTTAACAGCTACAGGCTACAGGCTACAGTGTGCGCTAG 5140
Qy 4801 GCCAAAGCCCCCCCCGCTCTGGAGCGTCAATGTGAAAGTGTGACTCGACTCAAGCCC 4860
Db 5141 GCCAAAGCCCCCCCCGCTCTGGAGCGTCAATGTGAAAGTGTGACTCGACTCAAGCCC 5200
Qy 4861 ACACCTGTGGGCCCAACACCTCTCTGTACCGCTTGGGCTGTGTTACCAAGAGGTCAAC 4920
Db 5201 ACACCTGTGGGCCCAACACCTCTCTGTACCGCTTGGGCTGTGTTACCAAGAGGTCAAC 5260

Qy 4921 CTCACACATCCCGTGAAGAAATACATCGCCACCTGTGATCAAGCCGACTTGGAGGTCAATG 4980
Db 5261 CTCACACATCCCGTGAAGAAATACATCGCCACCTGTGATCAAGCCGACTTGGAGGTCAATG 5320
Qy 4981 ACCAGCAATGGGCTTTGGCAGGGGAGTCTTGGCGGCGGTGCGCGGTATTTGCTGTGGCG 5040
Db 5321 ACCAGCAATGGGCTTTGGCAGGGGAGTCTTGGCGGCGGTGCGCGGTATTTGCTGTGGCG 5380
Qy 5041 ACCGGGTGTGTTGATCATCGCGCGCTTGGCATTTAACACAGGACCGCTGTTGGCGCG 5100
Db 5381 ACCGGGTGTGTTGATCATCGCGCGCTTGGCATTTAACACAGGACCGCTGTTGGCGCG 5440
Qy 5101 GACCAAGAGTCTCTATAGAGGCTTTGATGAGATGAGAAATGTGCTCTAAGCGGCT 5160
Db 5441 GACCAAGAGTCTCTATAGAGGCTTTGATGAGATGAGAAATGTGCTCTAAGCGGCT 5500
Qy 5161 CTCATTGAAAGAGGGCGAGCGGATAGCCGAGATGCTGAATCCAGATTCAGAGCTTATTTG 5220
Db 5501 CTCATTGAAAGAGGGCGAGCGGATAGCCGAGATGCTGAATCCAGATTCAGAGCTTATTTG 5560
Qy 5221 CAGCAAGCTTCCAAACAGCTCAAGCATTAACCCACTGTGAGGCTTATGAGCCCAAG 5280
Db 5561 CAGCAAGCTTCCAAACAGCTCAAGCATTAACCCACTGTGAGGCTTATGAGCCCAAG 5620
Qy 5281 GTAGAACAAATCTGGGCCAAACACATGTGAACTTTCATTAGCGGCATTCGAATCTTCGCA 5340
Db 5621 GTAGAACAAATCTGGGCCAAACACATGTGAACTTTCATTAGCGGCATTCGAATCTTCGCA 5680
Qy 5341 GGACTATCAACCTGCCAGGAAACCTGCAAGTACCTTCATATAGCGCTTCACTAGTGGCGCC 5400
Db 5681 GGACTATCAACCTGCCAGGAAACCTGCAAGTACCTTCATATAGCGCTTCACTAGTGGCGCC 5740
Qy 5401 CTCACAGTCCGCTGCAACAGCACCATATCTTCTGAACATTTTGGGGGGCTGGCTA 5460
Db 5741 CTCACAGTCCGCTGCAACAGCACCATATCTTCTGAACATTTTGGGGGGCTGGCTA 5800
Qy 5461 GCATCCCAAAATTTGCAACACCCCGGGGGCCACTGGCTTGTGATGAGTGGAG 5520
Db 5801 GCATCCCAAAATTTGCAACACCCCGGGGGCCACTGGCTTGTGATGAGTGGAG 5860
Qy 5521 GCTGCCGTAGGCGATTAAGCTTAAGTAAAGTGTCTAGTGAATCTTGGCAGGGTATGTG 5580
Db 5861 GCTGCCGTAGGCGATTAAGCTTAAGTAAAGTGTCTAGTGAATCTTGGCAGGGTATGTG 5920
Qy 5581 GCGGGCATTTTGGGGGCTCTCGTGGCATTAAGTATGATCTGTGGCAGAAAGCCCTCATG 5640
Db 5921 GCGGGCATTTTGGGGGCTCTCGTGGCATTAAGTATGATCTGTGGCAGAAAGCCCTCATG 5980
Qy 5641 GAGGATGTGCTCAACTTGTGCTGCTGGAATTTCTGTCTCGGGTGCCTTGTAGTGGAGTTC 5700
Db 5981 GAGGATGTGCTCAACTTGTGCTGCTGGAATTTCTGTCTCGGGTGCCTTGTAGTGGAGTTC 6040
Qy 5701 ATCTGGCGGCGCATTTCTGGCGGACACATGTGGACCCGGGGAAAGCCCGCTCAATGGATG 5760
Db 6041 ATCTGGCGGCGCATTTCTGGCGGACACATGTGGACCCGGGGAAAGCCCGCTCAATGGATG 6100
Qy 5761 AATTAAGCTAATGCTTGTGCTTCCAGAGAAATCAGTGCGCCCCACCCACTAGTGAAG 5820
Db 6101 AATTAAGCTAATGCTTGTGCTTCCAGAGAAATCAGTGCGCCCCACCCACTAGTGAAG 6160
Qy 5821 GAGTCGAGATGCTGCGAGCGTGTGACCAACTTACTTGTGCTTCCATTAACCAAGCTG 5880
Db 6161 GAGTCGAGATGCTGCGAGCGTGTGACCAACTTACTTGTGCTTCCATTAACCAAGCTG 6220
Qy 5881 CTCAGAAAGCTCCAACATGAGATTAAGAGGATGCGCCCATTCCTATGCGCGGCTGTGG 5940
Db 6221 CTCAGAAAGCTCCAACATGAGATTAAGAGGATGCGCCCATTCCTATGCGCGGCTGTGG 6280
Qy 5941 CTCGCGAGATGTGTGGGCTGGGTTTGACACATCTTAACAGACTTTTAAAAATTTGGCTGACC 6000
Db 6281 CTCGCGAGATGTGTGGGCTGGGTTTGACACATCTTAACAGACTTTTAAAAATTTGGCTGACC 6340

OY	6001	TCGAAATTTATTCCTCCAAAGATGCCCCGGGCTCCCCCTTTGTCTCTGTCAAAAGGGGTACAG	6060
Dp	6341	TCGAAATTTATTCCTCCAAAGATGCCCCGGGCTCCCCCTTTGTCTCTGTCAAAAGGGGTACAG	6400
OY	6061	GGCGTGTGGGCGGGCACTGGGATCATGACACACAGGTGTCTTTCGGGGGCCAATATCTCT	6120
Dp	6401	GGCGTGTGGGCGGGCACTGGGATCATGACACACAGGTGTCTTTCGGGGGCCAATATCTCT	6460
OY	6121	GGCAATGTCCGGCTTGGGCTCCATGAGATCACGGGGCTTAAGACTGTGCATGAATATCTGG	6180
Dp	6461	GGCAATGTCCGGCTTGGGCTCCATGAGATCACGGGGCTTAAGACTGTGCATGAATATCTGG	6520
OY	6181	CAGGGGACCTTTCCTATTCATATTTGTTTACACGAGGGGCCAGTGTGTGCGGAAACCCGGCGCA	6240
Dp	6521	CAGGGGACCTTTCCTATTCATATTTGTTTACACGAGGGGCCAGTGTGTGCGGAAACCCGGCGCA	6580
OY	6241	AACCTTAAAGTGTGGCATCTGAGAGGGTGGCGGCTTCAGAGTACGGGGAGTGAAGCAGACAC	6300
Dp	6581	AACCTTAAAGTGTGGCATCTGAGAGGGTGGCGGCTTCAGAGTACGGGGAGTGAAGCAGACAC	6640
OY	6301	GGGTCAATACCACTATCAATAACAGGACTCAACACTGATTAACCTTGAAGTCCCCCTGCCAATA	6360
Dp	6641	GGGTCAATACCACTATCAATAACAGGACTCAACACTGATTAACCTTGAAGTCCCCCTGCCAATA	6700
OY	6361	CCCTCTCCCGAGTTTCTTTTCTGGGTGAGCGGAGTGCAGATCATAGGTTTCCCCCACA	6420
Dp	6701	CCCTCTCCCGAGTTTCTTTTCTGGGTGAGCGGAGTGCAGATCATAGGTTTCCCCCACA	6760
OY	6421	CCGAAAGCCGTTTTTCCGGGATGAGGTGTCTGTGCGTGTGGGCGTTAATATCAATTTGTGTCT	6480
Dp	6761	CCGAAAGCCGTTTTTCCGGGATGAGGTGTCTGTGCGTGTGGGCGTTAATATCAATTTGTGTCT	6820
OY	6481	GGGTCCCGAGCTTCTCTTGGGACCCCTGAAACCCGACACAGACGTATGATGTCCATGCTAACA	6540
Dp	6821	GGGTCCCGAGCTTCTCTTGGGACCCCTGAAACCCGACACAGACGTATGATGTCCATGCTAACA	6880
OY	6541	GATCATCTCATATCAACGGCGGAGACTGACAGCGGGCGTTTGAAGCGGGGTCACCCCCA	6600
Dp	6881	GATCATCTCATATCAACGGCGGAGACTGACAGCGGGCGTTTGAAGCGGGGTCACCCCCA	6940
OY	6601	TCCGAGGCGAGCTCTCTGGGCGAGCGCAGCTATTCGGCACATGTGTGTGCGAGCCACTTGCACC	6660
Dp	6941	TCCGAGGCGAGCTCTCTGGGCGAGCGCAGCTATTCGGCACATGTGTGTGCGAGCCACTTGCACC	7000
OY	6661	ACCAGCGCAAAAGCCTATGATGTGGAACATGTGTGATGCTAACCCTTTCATGTGGGGGGAGAT	6720
Dp	7001	ACCAGCGCAAAAGCCTATGATGTGGAACATGTGTGATGCTAACCCTTTCATGTGGGGGGAGAT	7060
OY	6721	GTGACTGGGATGAGTCTGGGTCCAAAGTGTGTCTGTGACTCTCTGCACCCATGTGCT	6780
Dp	7061	GTGACTGGGATGAGTCTGGGTCCAAAGTGTGTCTGTGACTCTCTGCACCCATGTGCT	7120
OY	6781	GAAAGAAAGAGCGACCTTGAAGCTTTCGATTACATCAGATTCAGATGCTTCCCAAGAAAGAG	6840
Dp	7121	GAAAGAAAGAGCGACCTTGAAGCTTTCGATTACATCAGATTCAGATGCTTCCCAAGAAAGAG	7180
OY	6841	TTTCCCAACAGCTTATACCGGGCTGTGGGCAACGGCTGTATTAACAACCAACCGCTTGTGGAAATCG	6900
Dp	7181	TTTCCCAACAGCTTATACCGGGCTGTGGGCAACGGCTGTATTAACAACCAACCGCTTGTGGAAATCG	7240
OY	6901	TGGAAGAGGCGCAGATTACAAACCGGCGCACTGTTCGGGCTGTGTCTTCCCTCTCTAGG	6960
Dp	7241	TGGAAGAGGCGCAGATTACAAACCGGCGCACTGTTCGGGCTGTGTCTTCCCTCTCTAGG	7300
OY	6961	AAAAACCCCGACGCTCCCCCAAGAGGGGCGCGACAGTGGGCTCTAAGTGAAGACTCCATA	7020
Dp	7301	AAAAACCCCGACGCTCCCCCAAGAGGGGCGCGACAGTGGGCTCTAAGTGAAGACTCCATA	7360
OY	7021	GGAGATGCGCTTTCGAAACGCTGGCCATTAAAGTCCTTTGGCCAGGCCCCCCCCAAGGGGGAT	7080
Dp	7361	GGAGATGCGCTTTCGAAACGCTGGCCATTAAAGTCCTTTGGCCAGGCCCCCCCCAAGGGGGAT	7420
OY	7081	TCAGGCGCTTTTCACGAGGGGCGGGGCGCTGCGATTTCGGCACTCAGACGCTCTGATGAG	7140

Db	7421	TCAGGCGCTTTCACAGGGGGCGGGCGCTGCGGATTCGGGACGTGACGCGCTCTTATGAG	7480
Oy	7411	TTGGCGCTTTCGAGAGACAGGTTTCATCTCTTCCATGCGCCCGCTCGAGGGAGGCTTGGAA	7200
Db	7481	TTGGCGCTTTCGAGAGACAGGTTTCATCTCTTCCATGCGCCCGCTCGAGGGAGGAGCTTGGAA	7540
Oy	7201	GATTCAGACCTGAGAGCTTGAGCAGGTGAGAGCCCAACCCCGCCCGCAGGGGGGGGTGGCA	7260
Db	7541	GATTCAGACCTGAGAGCTTGAGCAGGTGAGAGCCCAACCCCGCCCGCAGGGGGGGGTGGCA	7600
Oy	7261	GCTCCGGGCTCGAGCTCGGGGCTCTGGTCTACCTTGGCTCCGAGGAGGACGACTCCGTCGG	7320
Db	7601	GCTCCGGGCTCGAGCTCGGGGCTCTGGTCTACCTTGGCTCCGAGGAGGACGACTCCGTCGG	7660
Oy	7321	TGCTGCTCCATGTCATCTACTCTTGAGACCGGGGCTCTAATATCTCTTGTATGTCGCGAAG	7380
Db	7661	TGCTGCTCCATGTCATCTACTCTTGAGACCGGGGCTCTAATATCTCTTGTATGTCGCGAAG	7720
Oy	7381	GAGAAATTACCGATTAAACCCCTTGAGCAACTCCCTGTGCGATATCAACAAGGTGAC	7440
Db	7721	GAGAAATTACCGATTAAACCCCTTGAGCAACTCCCTGTGCGATATCAACAAGGTGAC	7780
Oy	7441	TGTATCCACAACAAAGAGCGCTTACATTAAGGCGCTAAATAGTAACTTTTGTATGATGATGCA	7500
Db	7781	TGTATCCACAACAAAGAGCGCTTACATTAAGGCGCTAAATAGTAACTTTTGTATGATGATGCA	7840
Oy	7501	GTGCTCGACTCTACTACGACTCAGTCTTAAAGGACATTAAGCTAGAGCGGCTCCAAAGTC	7560
Db	7841	GTGCTCGACTCTACTACGACTCAGTCTTAAAGGACATTAAGCTAGAGCGGCTCCAAAGTC	7900
Oy	7561	ACCGCAAGGCTCCTCAACATGAGAGAGGCTTGCGATTAAACCCACCCCATTTCTGCAAGA	7620
Db	7901	ACCGCAAGGCTCCTCAACATGAGAGAGGCTTGCGATTAAACCCACCCCATTTCTGCAAGA	7960
Oy	7621	TCTAATATAGGCTTTGGGGCTTAAGAGAGTCCGACGTTGTCCGGAGGGCGCTTAACAC	7680
Db	7961	TCTAATATAGGCTTTGGGGCTTAAGAGAGTCCGACGTTGTCCGGAGGGCGCTTAACAC	8020
Oy	7681	ATCAAGTCGCGTGGAAAGGACCTCCGAGAGACTCAAGAAACCAATTCGCCACAACATTT	7740
Db	8021	ATCAAGTCGCGTGGAAAGGACCTCCGAGAGACTCAAGAAACCAATTCGCCACAACATTT	8080
Oy	7741	ATGCGCAAAATATAGAGTGTCTGCGTGGAGCCCGCACCAAGGGGGGCGAAGAAACAGCTGCG	7800
Db	8081	ATGCGCAAAATATAGAGTGTCTGCGTGGAGCCCGCACCAAGGGGGGCGAAGAAACAGCTGCG	8140
Oy	7801	CTTATCGTTTAAACCTTGACCTCGGCGTCAAGGCTCTGCGAGAGATGCGCTTTATGACATT	7860
Db	8141	CTTATCGTTTAAACCTTGACCTCGGCGTCAAGGCTCTGCGAGAGATGCGCTTTATGACATT	8200
Oy	7861	ACACAAAAACCTTCTCAGCGGGGTGATGGGGGCTTTATGAAATTCAGATTTCCCGCGCT	7920
Db	8201	ACACAAAAACCTTCTCAGCGGGGTGATGGGGGCTTTATGAAATTCAGATTTCCCGCGCT	8260
Oy	7921	CAGCGGGTATAGTTTCTCTTGAAGCATGGGCGGAAAGAAAGACCTTATGGGTTTTTCG	7980
Db	8261	CAGCGGGTATAGTTTCTCTTGAAGCATGGGCGGAAAGAAAGACCTTATGGGTTTTTCG	8320
Oy	7981	TATGATATCCCGATGCTTTTGACTCAACCGTCACTGAGAGACATCGACTGAGAGTTC	8040
Db	8321	TATGATATCCCGATGCTTTTGACTCAACCGTCACTGAGAGACATCGACTGAGAGTTC	8380
Oy	8041	ATATATGGGGCGTGCCTTTCGCGGAGAGGCGCCACACTGCGCATCACTCGGTAATGAG	8100
Db	8381	ATATATGGGGCGTGCCTTTCGCGGAGAGGCGCCACACTGCGCATCACTCGGTAATGAG	8440
Oy	8101	AGACTTTACCTGGAGAGGCGCTATGTTCAACAGCAAGGCGCAAACTGCGGATACAGGCGT	8160
Db	8441	AGACTTTACCTGGAGAGGCGCTATGTTCAACAGCAAGGCGCAAACTGCGGATACAGGCGT	8500
Oy	8161	TGCGCGCCAGCGGGGTGCTCAACATGACATGGGAGACCATATCATGTCTACTGAAA	8220

OY	841	TCGGGTGGGGGATGCTCCGAGCCCAATGTTCAATGTCCTGCGCGAGACACACTGGCTTT	900
Db	1181	TGCGGTGGGGGATGCTCCGAGCCCAATGTTCAATGTCCTGCGCGAGACACACTGGCTTT	1240
OY	901	GTCCAAAGCTCAATATGGCTCCATCTACCTCGTATCAATCACTGGAACACCGCATGGACATGG	960
Db	1241	GTCCAAAGCTCAATATGGCTCCATCTACCTCGTATCAATCACTGGAACACCGCATGGACATGG	1300
OY	961	GACATGATGATGAATCTGGTGGCCCAACGGCTACCATGATCTTGGCGTACGGATGCTGTCTC	1020
Db	1301	GACATGATGATGAATCTGGTGGCCCAACGGCTACCATGATCTTGGCGTACGGATGCTGTCTC	1360
OY	1021	CCCGAGGTCATTAATAGATCATTAATAGCGGGGCTCATTTGGGGCGTCAATGTTCCGCTTGGCC	1080
Db	1361	CCCGAGGTCATTAATAGATCATTAATAGCGGGGCTCATTTGGGGCGTCAATGTTCCGCTTGGCC	1420
OY	1081	TACTTCTTATAGCAGGAGCGTGGGCGAAAGTGCATTGTCATCTCTTGTGGCGCGGG	1140
Db	1421	TACTTCTTATAGCAGGAGCGTGGGCGAAAGTGCATTGTCATCTCTTGTGGCGCGGG	1480
OY	1141	GTGACGCGGCGACCCCATATCTGTTGGGGGTTCTGCGCGCAGACACCGGGCGCTCAAC	1200
Db	1481	GTGACGCGGCGACCCCATATCTGTTGGGGGTTCTGCGCGCAGACACCGGGCGCTCAAC	1540
OY	1201	AGCTTATTTGACATGGGGCCCAAGGACAAAATCCAGCTCGTTTAAACCAATATGGACGTGG	1260
Db	1541	AGCTTATTTGACATGGGGCCCAAGGACAAAATCCAGCTCGTTTAAACCAATATGGACGTGG	1600
OY	1261	CACATCAACCGGACCGGCTGTAGCTGCAATGATCTCTTGGACACCGGCTTTATGCGGCT	1320
Db	1601	CACATCAACCGGACCGGCTGTAGCTGCAATGATCTCTTGGACACCGGCTTTATGCGGCT	1660
OY	1321	CTGTTTCAACACCCACAGCTTCAACTCGTCAAGATATGCCGAAGCATATGCCCTGCGCGC	1380
Db	1661	CTGTTTCAACACCCACAGCTTCAACTCGTCAAGATATGCCGAAGCATATGCCCTGCGCGC	1720
OY	1381	AGTATCGAGGCTTCCGGGTGGGATGGGGCGCTTGGCAATATGAGATTAATGTCAACAAT	1440
Db	1721	AGTATCGAGGCTTCCGGGTGGGATGGGGCGCTTGGCAATATGAGATTAATGTCAACAAT	1780
OY	1441	CCAGAGGATATGAGACCCCTATTTGCTGGGACATACCCACCAAGGACGTGGGTGGTCTCC	1500
Db	1781	CCAGAGGATATGAGACCCCTATTTGCTGGGACATACCCACCAAGGACGTGGGTGGTCTCC	1840
OY	1501	GCGAAGACTGTGTGTGGCCCACTGTATCTGTTTCAACCCCAAGCCAGTGTATGGGACAG	1560
Db	1841	GCGAAGACTGTGTGTGGCCCACTGTATCTGTTTCAACCCCAAGCCAGTGTATGGGACAG	1900
OY	1561	ACCGACAGGCTTGGAGCGGCCCATTTACAGTGGGGGAGAAATGAGACAGATGTCTTCTA	1620
Db	1901	ACCGACAGGCTTGGAGCGGCCCATTTACAGTGGGGGAGAAATGAGACAGATGTCTTCTA	1960
OY	1621	TTTGAACAGCACTTCGACCCACCGCTGGGGTCAATGTTTCGGCTGCACTGTGATGAATCTTCT	1680
Db	1961	TTTGAACAGCACTTCGACCCACCGCTGGGGTCAATGTTTCGGCTGCACTGTGATGAATCTTCT	2020
OY	1681	GGCTTACACCAAGACTTGGGCGGCAACACCTCGCGCTACTAGAGCTGACTTCAAGCGGACG	1740
Db	2021	GGCTTACACCAAGACTTGGGCGGCAACACCTCGCGCTACTAGAGCTGACTTCAAGCGGACG	2080
OY	1741	ACGGAAGCTGTGTGTGCCCAACGGGACTGTTTATGAGAGACATCTGTATACCACTTACCTCAAA	1800
Db	2081	ACGGAAGCTGTGTGTGCCCAACGGGACTGTTTATGAGAGACATCTGTATACCACTTACCTCAAA	2140
OY	1801	TGCGGCTCTGGGGCCCTGGCTCAAGCCAAAGTGGCTGATTCGACTACCCCTACAGGCTCTGG	1860
Db	2141	TGCGGCTCTGGGGCCCTGGCTCAAGCCAAAGTGGCTGATTCGACTACCCCTACAGGCTCTGG	2200
OY	1861	CATTACCCCTGACAGTTAACTTATACATCTTCAAAAAATAGAGTATATGTGGAGGGGGTT	1920
Db	2201	CATTACCCCTGACAGTTAACTTATACATCTTCAAAAAATAGAGTATATGTGGAGGGGGTT	2260
OY	1921	GAGCACAAGGCTCACGGCTGATGCAATTTTCACTCGTGGGAGTCTGTGCACCTTGAAGGAC	1980

Db	2261	GAGACAGGGCTCACGGGTGCATGCAAATTTTCACTCGTGGGGAATCGTTGGCAACTTGGAGGAC	2320
OY	1981	AGAGACAGAGGTAAACATGTCTCTCTTTGGTGCACCTCCACGAGGAATGGGCATTTTAACT	2040
Db	2321	AGAGACAGAGGTCAACATGTCTCTCTTTGGTGCACCTCCACGAGGAATGGGCATTTTAACT	2380
OY	2041	TGCTCTTACTCGGACCTGCCCGCTTGTGACTGTGTCTTCTCCACCTCCACCAAAATC	2100
Db	2381	TGCTCTTACTCGGACCTGCCCGCTTGTGACTGTGTCTTCTCCACCTCCACCAAAATC	2440
OY	2101	GTGGACGTAAATTCAATGTATGGCCATATACCTGGCCCTCAAAAATACATGCTGCGATGG	2160
Db	2441	GTGGACGTAAATTCAATGTATGGCCATATACCTGGCCCTCAAAAATACATGCTGCGATGG	2500
OY	2161	GAGTGGGTAAATACTCTTAATTCCTGCTCTTAAGCGAGCGCAGAGGTTTGCCTCGTAAATG	2220
Db	2501	GAGTGGGTAAATACTCTTAATTCCTGCTCTTAAGCGAGCGCAGAGGTTTGCCTCGTAAATG	2560
OY	2221	ATGCTCATCTTGTGTGGCCAGGCGGAGAGCAGCACTAGAGAACTGGTCACTTTTGAACGCT	2280
Db	2561	ATGCTCATCTTGTGTGGCCAGGCGGAGAGCAGCACTAGAGAACTGGTCACTTTTGAACGCT	2620
OY	2281	GCGAGGCCACTTACGTGCATATGGCTTCCATATTTTGTGATCTTTTGTGGCTGCTTGG	2340
Db	2621	GCGAGGCCACTTACGTGCATATGGCTTCCATATTTTGTGATCTTTTGTGGCTGCTTGG	2680
OY	2341	TACATCAAGGGGTGGGGATGACCCCTTAGCTACCTATTCCTCACTGAGGCTGTGGTCTT	2400
Db	2681	TACATCAAGGGGTGGGGATGACCCCTTAGCTACCTATTCCTCACTGAGGCTGTGGTCTT	2740
OY	2401	AGCCTCATGCTCTAGCATTTGCCCAACAGGCTTATGCTTATGACGCATCTGTGCATGGC	2460
Db	2741	AGCCTCATGCTCTAGCATTTGCCCAACAGGCTTATGCTTATGACGCATCTGTGCATGGC	2800
OY	2461	CAGATAGGAGCGGCTCTGCTGGTATGATCACTCTTTACTCTCACCCCGGGTATTAAG	2520
Db	2801	CAGATAGGAGCGGCTCTGCTGGTATGATCACTCTTTACTCTCACCCCGGGTATTAAG	2860
OY	2521	ACCCCTTCAGCGGGTTTTGTGATGGTGTGCTATCTTCTGACCTGGGGGAGCTAATG	2580
Db	2861	ACCCCTTCAGCGGGTTTTGTGATGGTGTGCTATCTTCTGACCTGGGGGAGCTAATG	2920
OY	2581	GTCGAGGAGTGGGACCACTATGACGATGCGCGGTGGCGCGTATGCACTATATGGGCG	2640
Db	2921	GTCGAGGAGTGGGACCACTATGACGATGCGCGGTGGCGCGTATGCACTATATGGGCG	2980
OY	2641	GTGGCCATATTTCAACCAAGGTGTGTGTTGACATPACCAAGTGGCTTTGGCGGTGCTT	2700
Db	2981	GTGGCCATATTTCAACCAAGGTGTGTGTTGACATPACCAAGTGGCTTTGGCGGTGCTT	3040
OY	2701	GGGCTGCTTAACCTCTCAAAAGGTGCTTTGAAGCGCGGTGCGCTAATCTGTCAGGGCTAC	2760
Db	3041	GGGCTGCTTAACCTCTCAAAAGGTGCTTTGAAGCGCGGTGCGCTAATCTGTCAGGGCTAC	3100
OY	2761	GCTCTACTGAGATGTGCAACATGCAAGGCACTTCGCGGGGGGAGGATAGCTCCAGATG	2820
Db	3101	GCTCTACTGAGATGTGCAACATGCAAGGCACTTCGCGGGGGGAGGATAGCTCCAGATG	3160
OY	2821	GCGCTACTAGCCCTTGGCAAGTGAATGSCACTTATCATATGACCACTCAACCCCTATG	2880
Db	3161	GCGCTACTAGCCCTTGGCAAGTGAATGSCACTTATCATATGACCACTCAACCCCTATG	3220
OY	2881	TTCGGATTGGGCTCTAGTGTGGCCGCGGAGACCTGGCGGTGCGCGTTAGAGCTTACATCTTC	2940
Db	3221	TTCGGATTGGGCTCTAGTGTGGCCGCGGAGACCTGGCGGTGCGCGTTAGAGCTTACATCTTC	3280
OY	2941	AGTCCGATGAGAAAGATCATATTCTGTGGGAGCGGAGACAGCTGCTTGTGGGACATTT	3000
Db	3281	AGTCCGATGAGAAAGATCATATTCTGTGGGAGCGGAGACAGCTGCTTGTGGGACATTT	3340
OY	3001	TTTACAGGACTTCCCGTGTCCGCGGACCTTGGTCCGAGAGTCTCTCTTGGCCACTGAT	3060

Db 3341 TTACACGACCTTCCGTCGCGCCGACCTTGTCGGAGATCTCTTGTGCGCCAGCTGAT 3400
Oy 3061 GGCTATACCTCCAAAGGGGTGAGTCTTCTGGCCCCATCACTGGCTTAAAGCCACAGACA 3120
Db 3401 GGCTATACCTCCAAAGGGGTGAGTCTTCTGGCCCCATCACTGGCTTAAAGCCACAGACA 3460
Oy 3121 CGTGGCCCTTTTGGGCAACCATAGTGTGAGCATGACGGGGCCGCAACAAGACAAGAGCT 3180
Db 3461 CGTGGCCCTTTTGGGCAACCATAGTGTGAGCATGACGGGGCCGCAACAAGACAAGAGCT 3520
Oy 3181 GGGGAAATTCAGAGTCTCTTCCACAGTCACTAGTCTCTTCCGAAACATCCATCTCGGG 3240
Db 3521 GGGGAAATTCAGAGTCTCTTCCACAGTCACTAGTCTCTTCCGAAACATCCATCTCGGG 3580
Oy 3241 GTTTGTGAGCTGTCTACATGAGCTGGCAACAAGACTTGGCCGGCTCAAGGGATCCG 3300
Db 3581 GTTTGTGAGCTGTCTACATGAGCTGGCAACAAGACTTGGCCGGCTCAAGGGATCCG 3640
Oy 3301 GTCAAGCAGATGTACTCCAGTGTCTGAGGGGGAATTAGTAGGTGGCCCAAGCCCTGGG 3360
Db 3641 GTCAAGCAGATGTACTCCAGTGTCTGAGGGGGAATTAGTAGGTGGCCCAAGCCCTGGG 3700
Oy 3361 ACTAAATCTTTGAGAGCCGTGACAGTGTGAGCCGCTGACCTGTATCTGTGTACCGCGAAC 3420
Db 3701 ACTAAATCTTTGAGAGCCGTGACAGTGTGAGCCGCTGACCTGTATCTGTGTACCGGAAAC 3760
Oy 3421 GCTGATGTCACTCCGGCTCGAAGACGCGGGGCAAAACGGGGAGCGCTACTCTCCCGAGA 3480
Db 3761 GCTGATGTCACTCCGGCTCGAAGACGCGGGGCAAAACGGGGAGCGCTACTCTCCCGAGA 3820
Oy 3481 CCTCTTTTCACCTTGAAGGGGTCTCTCAGAGAGCGCCGATGTATGCCCCAGGGGACACGCT 3540
Db 3821 CCTCTTTTCACCTTGAAGGGGTCTCTCAGAGAGCGCCGATGTATGCCCCAGGGGACACGCT 3880
Oy 3541 GTCCGAGCTTCCCGGGAGAGCTGTGTCTCGGGCCGCTGTAAGTCCATAGATTTATC 3600
Db 3881 GTCCGAGCTTCCCGGGAGAGCTGTGTCTCGGGCCGCTGTAAGTCCATAGATTTATC 3940
Oy 3601 CCCGTTGAGACACTGAGCATCGTCAAGCGGTCCCCCACTTTAGTGAACAAGACACA 3660
Db 3941 CCCGTTGAGACACTGAGCATCGTCAAGCGGTCCCCCACTTTAGTGAACAAGACACA 4000
Oy 3661 CCTGTGTGCCCCAGACCTATCAGGTGGGGTAACTTTCATGCCCCGACCTGGAGTGAAG 3720
Db 4001 CCTGTGTGCCCCAGACCTATCAGGTGGGGTAACTTTCATGCCCCGACCTGGAGTGAAG 4060
Oy 3721 AGCACAAGATTCCTGTGAGCATATGCTCTCAAGGGGTAAAGTGTAGTCTTAATCCC 3780
Db 4061 AGCACAAGATTCCTGTGAGCATATGCTCTCAAGGGGTAAAGTGTAGTCTTAATCCC 4120
Oy 3781 TCAGTGTGCGCACCTCGGGGTTTGGGGCTGTAATTGTCTAAGGCACATGATCAATCCC 3840
Db 4121 TCAGTGTGCGCACCTCGGGGTTTGGGGCTGTAATTGTGTAAAGGCACATGATCAATCCC 4180
Oy 3841 AACATTAGAGACTGAGTGAAGACTGTGACGACCGGGGCGCCATCACTGATCCACATAT 3900
Db 4181 AACATTAGAGACTGAGTGAAGACTGTGACGACCGGGGCGCCATCACTGATCCACATAT 4240
Oy 3901 GGGCAATTCCTCGCGAGTGGGGCTGTGCGGGCCGCTACGACATATATATATGAT 3960
Db 4241 GGGCAATTCCTCGCGAGTGGGGCTGTGCGGGCCGCTACGACATATATATATGAT 4300
Oy 3961 GAATGCCATGCCGTGAGACTCTACACACATCTTGGCATCGAAACAGTCTTGATCAAGCA 4020
Db 4301 GAATGCCATGCCGTGAGACTCTACACACATCTTGGCATCGAAACAGTCTTGATCAAGCA 4360
Oy 4021 GAGACAGCTGGGTGAGACTTATCTGTGCTGCTACAGCTACGCCCTTGAGTCAATGACA 4080
Db 4361 GAGACAGCTGGGTGAGACTTATCTGTGCTGCTACAGCTACGCCCTTGAGTCAATGACA 4420
Oy 4081 ACCCCCCACCCCAATAGAGAGGTGGCCCTTGGGCGAGAGAGGAGATCCCTTCTAT 4140
Db 4421 ACCCCCCACCCCAATAGAGAGGTGGCCCTTGGGCGAGAGAGGAGATCCCTTCTAT 4480

Oy 4141 GGGAGGGCGATTCCCGTCTTACATCAAGGAGGAAAGACATGTGATCTTCCGCAATTCA 4200
Db 4481 GGGAGGGCGATTCCCGTCTTACATCAAGGAGGAAAGACATGTGATCTTCCGCAATTCA 4540
Oy 4201 AAGAAAAAGTGTGACGAGCTCGCGGCGCCCTTTCGGGGTATGAGGCTTGAACCTCAGTGCCA 4260
Db 4541 AAGAAAAAGTGTGACGAGCTCGCGGCGCCCTTTCGGGGTATGAGGCTTGAACCTCAGTGCCA 4600
Oy 4261 TACTACAGAGGGTGTGACGCTCCGTAATACCACTCAGGGAGAGGTAGTGGTCCGCGC 4320
Db 4601 TACTACAGAGGGTGTGACGCTCCGTAATACCACTCAGGGAGAGGTAGTGGTCCGCGC 4660
Oy 4321 ACCGAGCCCTCATGACAGGGTATACTGGGGAATTGACTCCGTGATCGACTGCAACGTA 4380
Db 4661 ACCGAGCCCTCATGACAGGGTATACTGGGGAATTGACTCCGTGATCGACTGCAACGTA 4720
Oy 4381 GCGGTCACTCAAGTTGTAGCTTCAAGTTTAAAGCCCACTTCACTCAACACAGATT 4440
Db 4721 GCGGTCACTCAAGTTGTAGCTTCAAGTTTAAAGCCCACTTCACTCAACACAGATT 4780
Oy 4441 GTCCCTCAAGAGCGCTGTCAAGTACGACCGCCGGGGTCCGACGGGTAGGGGAAAGCTG 4500
Db 4781 GTCCCTCAAGAGCGCTGTGTCAAGTACGACCGCCGGGGTCCGACGGGTAGGGGAAAGCTG 4840
Oy 4501 GGCATTTATAGTATGTTTCACTGGTGAAGCGAGCTCAGGAATGTTTGAAGTGTAGTG 4560
Db 4841 GGCATTTATAGTATGTTTCACTGGTGAAGCGAGCTCAGGAATGTTTGAAGTGTAGTG 4900
Oy 4561 CTCTGTAGTGTCTACGACGAGGGGCCGATGATAGAGTCAACACATCGAGACACACC 4620
Db 4901 CTCTGTAGTGTCTACGACGAGGGGCCGATGATAGAGTCAACACATCGAGACACACC 4960
Oy 4621 GTCAGGCTCAGGGCGATTTCAACACGCGCGGTTGCTGTGTGCCAAGACCATCTTGAAG 4680
Db 4961 GTCAGGCTCAGGGCGATTTCAACACGCGCGGTTGCTGTGTGCCAAGACCATCTTGAAG 5020
Oy 4681 TTTTGGAGGCACTTTTCAACCGGCTCACAACATAGATGCCACTTCTTTCCAAACA 4740
Db 5021 TTTTGGAGGCACTTTTCAACCGGCTCACAACATAGATGCCACTTCTTTCCAAACA 5080
Oy 4741 AAGCAATCGGGGGAATTTCCGATCTTAACGCTTACAGGCTACAGTGTGCGCTAGG 4800
Db 5081 AAGCAATCGGGGGAATTTCCGATCTTAACGCTTACAGGCTACAGTGTGCGCTAGG 5140
Oy 4801 GCCAAGCCCCCCCCCGTCTGTGGGAGTGAATGGAAGTGTGAAGTCAAGGCC 4860
Db 5141 GCCAAGCCCCCCCCCGTCTGTGGGAGTGAATGGAAGTGTGAAGTCAAGGCC 5200
Oy 4861 ACACTGCTGGCCCCCAACACTCTCTGTACCGGCTTGGGCTGTGTACCAACGAGTCAAC 4920
Db 5201 ACACTGCTGGCCCCCAACACTCTCTGTACCGGCTTGGGCTGTGTACCAACGAGTCAAC 5260
Oy 4921 CTCAACATCCCGTGAAGAAATACATCGGCCACTGTCATGCAAGCCGACTTGAAGTCAAG 4980
Db 5261 CTCAACATCCCGTGAAGAAATACATCGGCCACTGTCATGCAAGCCGACTTGAAGTCAAG 5320
Oy 4981 ACCAGCAGATGGGCTTGGCAGGGGAGTCTTGGCGGCGTCCGCGGATTCCTGAGG 5040
Db 5321 ACCAGCAGATGGGCTTGGCAGGGGAGTCTTGGCGGCGTCCGCGGATTCCTGAGG 5380
Oy 5041 ACCGGGTGTGTCATCATCGGCGCTTGCACATTAACCAAGGAGCCGTGTTGCGCG 5100
Db 5381 ACCGGGTGTGTCATCATCGGCGCTTGCACATTAACCAAGGAGCCGTGTTGCGCG 5440
Oy 5101 GACAAAGAGGTCTCTATAGAGCTTTTGAATGAGATGAGGAATGTGCTTAAAGCGGCT 5160
Db 5441 GACAAAGAGGTCTCTATAGAGCTTTTGAATGAGATGAGGAATGTGCTTAAAGCGGCT 5500
Oy 5161 CTCAATGAAGAGGGGAGCGAGTACCGAGATGCTGAAGTCCAAAGTCCAAAGCTTATTTG 5220
Db 5501 CTCAATGAAGAGGGGAGCGAGTACCGAGATGCTGAAGTCCAAAGTCCAAAGCTTATTTG 5560

QY 5221 CAGCAAGCTTCAAAACAGCTCAAGACATACAAACCACTGTGAGGCTTCATGGCCCAAG 5280
| | | | |
Db 5561 CACCAAGCTTCCAAACAGCTCAAGACATACAAACCACTGTGAGGCTTCATGGCCCAAG 5620
| | | | |
QY 5281 GTTGAACATTTGGGGCAAAACACATGTGAACCTTCATTAGCGGCATCCAAATCTTCGA 5340
| | | | |
Db 5621 GTTGAACATTTGGGGCAAAACACATGTGAACCTTCATTAGCGGCATCCAAATCTTCGA 5680
| | | | |
QY 5341 GGACTATCAACACATGGCAGGAAACCCCTGAGTAGCTTCATGATGGGCTTCAGTGGCCG 5400
| | | | |
Db 5681 GGACTATCAACACATGGCAGGAAACCCCTGAGTAGCTTCATGATGGGCTTCAGTGGCCG 5740
| | | | |
QY 5401 CTCACAGTCCGCTGTCAACAGACCACTATCTTCTCAAACTTTTGGGGGCTGGCTA 5460
| | | | |
Db 5741 CTCACAGTCCGCTGTCAACAGACCACTATCTTCTCAAACTTTTGGGGGCTGGCTA 5800
| | | | |
QY 5461 GCATCCCAATTGACCAACCCGCGGGGCCACTGGCTTTCGTTGCACTGAGCTTAAGTGA 5520
| | | | |
Db 5801 GCATCCCAATTGACCAACCCGCGGGGCCACTGGCTTTCGTTGCACTGAGCTTAAGTGA 5860
| | | | |
QY 5521 GCTGCCGTAGGAGATATAGGCTTAAGTGAAGTGTATGAGACATCTCGGAGAGGTAAGT 5580
| | | | |
Db 5861 GCTGCCGTAGGAGATATAGGCTTAAGTGAAGTGTATGAGACATCTCGGAGAGGTAAGT 5920
| | | | |
QY 5581 GCGGGCATTTGCGGGGCTCTGTCGCAATTCAAGATCATGTCTGGCAGAAACCTTCATG 5640
| | | | |
Db 5921 GCGGGCATTTGCGGGGCTCTGTCGCAATTCAAGATCATGTCTGGCAGAAACCTTCATG 5980
| | | | |
QY 5641 GAGGATGTGCTCAACTTGTGCTGGAATTTCTGCTCGGGGTGCTTTGGTAGTGAAGTGC 5700
| | | | |
Db 5981 GAGGATGTGCTCAACTTGTGCTGGAATTTCTGCTCGGGGTGCTTTGGTAGTGAAGTGC 6040
| | | | |
QY 5701 ATCTGGCGCGGCAATTTCGCGCCGACACATGTGAACCGGGGGAAGCGCGCTCAATGATG 5760
| | | | |
Db 6041 ATCTGGCGCGGCAATTTCGCGCCGACACATGTGAACCGGGGGAAGCGCGCTCAATGATG 6100
| | | | |
QY 5761 AATAGACTCATTTGCTTTCAGAGAAATCACTCGCCCAACCACTACATGATGAGC 5820
| | | | |
Db 6101 AATAGACTCATTTGCTTTCAGAGAAATCACTCGCCCAACCACTACATGATGAGC 6160
| | | | |
QY 5821 GAGTCGAGTGCCTGCGAGCGTGTGACCACTAATTGGCTTCCCTTAACATACCAAGCTG 5880
| | | | |
Db 6161 GAGTCGAGTGCCTGCGAGCGTGTGACCACTAATTGGCTTCCCTTAACATACCAAGCTG 6220
| | | | |
QY 5881 CTGAGAAAGACTCCAAACATGGAATTAATCTAGAGACTGCCCCATCCATGCGCGCTGTG 5940
| | | | |
Db 6221 CTGAGAAAGACTCCAAACATGGAATTAATCTAGAGACTGCCCCATCCATGCGCGCTGTG 6280
| | | | |
QY 5941 CTCGCGCATGTGTGGAGCTGGATTTTGACACATCTTAACAGACTTTAAATAATGGCTGAC 6000
| | | | |
Db 6281 CTCGCGCATGTGTGGAGCTGGATTTTGACACATCTTAACAGACTTTAAATAATGGCTGAC 6340
| | | | |
QY 6001 TCCAAATTATTTCCAAAGATGCGCGCTTCCTTTGTCTCTGTCAAAAAGGGGTACAG 6060
| | | | |
Db 6341 TCCAAATTATTTCCAAAGATGCGCGCTTCCTTTGTCTCTGTCAAAAAGGGGTACAG 6400
| | | | |
QY 6061 GGGGTGTGGGCGGCACTGGCATCATAGACACACAGGTGTCTTGTGCGCGCATATCTCT 6120
| | | | |
Db 6401 GGGGTGTGGGCGGCACTGGCATCATAGACACACAGGTGTCTTGTGCGCGCATATCTCT 6460
| | | | |
QY 6121 GGCATGTCCGCTTGGGCTCATAGAGAAATCAGGGGCTTAAGACTGTGAGAAATATCTGG 6180
| | | | |
Db 6461 GGCATGTCCGCTTGGGCTCATAGAGAAATCAGGGGCTTAAGACTGTGAGAAATATCTGG 6520
| | | | |
QY 6181 CAGGGGACCTTTCTCATATTTGTATACAGAGGGGCAATGCTGCCGAAACCCGCGCA 6240
| | | | |
Db 6521 CAGGGGACCTTTCTCATATTTGTATACAGAGGGGCAATGCTGCCGAAACCCGCGCA 6580
| | | | |
QY 6241 AACTTTAAGGTGCGCATCTGAGAGGTGCGCGCTCAAGATACGCGAGGTGACGCAAGAC 6300
| | | | |
Db 6581 AACTTTAAGGTGCGCATCTGAGAGGTGCGCGCTCAAGATACGCGAGGTGACGCAAGAC 6640
| | | | |
QY 6301 GGGTCAATACCACTTAACAGGACTCACCACTGATTAATTGAAAGTCCCCCTGCACTA 6360
| | | | |

Db 6641 GGGTCAATACCACTTAACAGGACTCACCACTGATTAATTGAAAGTCCCCCTGCACTA 6700
| | | | |
QY 6361 CCTCTCCGAGTTCCTTTCTTGGGTGAGCGAGTGCATATAGTGTGCCCCACA 6420
| | | | |
Db 6701 CCTCTCCGAGTTCCTTTCTTGGGTGAGCGAGTGCATATAGTGTGCCCCACA 6760
| | | | |
QY 6421 CCGAAGCCGTTTTCCGGGAGTAGGTCTGTCTGGGTGGGCTTAATTGATTTGTGTGC 6480
| | | | |
Db 6761 CCGAAGCCGTTTTCCGGGAGTAGGTCTGTCTGGGTGGGCTTAATTGATTTGTGTGC 6820
| | | | |
QY 6481 GGGTCCAGCTTCCTTTCGACCTTGAACCCGACACAGACGTATATGTCATGCTAACA 6540
| | | | |
Db 6821 GGGTCCAGCTTCCTTTCGACCTTGAACCCGACACAGACGTATATGTCATGCTAACA 6880
| | | | |
QY 6541 GATTCATCTCATATACAGCGGAGACTGACAGCCGCGCTTACGCGGGGCTACCCCA 6600
| | | | |
Db 6881 GATTCATCTCATATACAGCGGAGACTGACAGCCGCGCTTACGCGGGGCTACCCCA 6940
| | | | |
QY 6601 TCCGAGGCAAGCTCTCGGAGAGCTATGGGCAACATGCTGGGAGGCACTGACCC 6660
| | | | |
Db 6941 TCCGAGGCAAGCTCTCGGAGAGCTATGGGCAACATGCTGGGAGGCACTGACCC 7000
| | | | |
QY 6661 ACCCAGGCAAGCTCTATGATGTGACATGTGGAATGCTTAACCTGTTCAATGGGGCGAT 6720
| | | | |
Db 7001 ACCCAGGCAAGCTCTATGATGTGACATGTGGAATGCTTAACCTGTTCAATGGGGCGAT 7060
| | | | |
QY 6721 GTGACTCGATATAGTCTGGGTCCAAAGTGTGTTCTGACCTCTCCACCCATGCTC 6780
| | | | |
Db 7061 GTGACTCGATATAGTCTGGGTCCAAAGTGTGTTCTGACCTCTCCACCCATGCTC 7120
| | | | |
QY 6781 GAGGAAAGGAGGAGCACTTGAAGCTTGCATTCAGAAATATAGTCTCCCAAGAGAG 6840
| | | | |
Db 7121 GAGGAAAGGAGGAGCACTTGAAGCTTGCATTCAGAAATATAGTCTCCCAAGAGAG 7180
| | | | |
QY 6841 TTCCACACAGCTTTTACCGGCTTGGGACCGGCTGATTAACAACCAACGCTTGTGAATG 6900
| | | | |
Db 7181 TTCCACACAGCTTTTACCGGCTTGGGACCGGCTGATTAACAACCAACGCTTGTGAATG 7240
| | | | |
QY 6901 TGGAAAAGGCGAGATTAACCAACCGGCACTGTTGCGGGCTGTGCTCTCTCTCTAGG 6960
| | | | |
Db 7241 TGGAAAAGGCGAGATTAACCAACCGGCACTGTTGCGGGCTGTGCTCTCTCTCTAGG 7300
| | | | |
QY 6961 AAAACCCGACGCTCCGCAAGAGGCGCGGACAGTGGGCTTAAGTGAAGCTCATTA 7020
| | | | |
Db 7301 AAAACCCGACGCTCCGCAAGAGGCGCGGACAGTGGGCTTAAGTGAAGCTCATTA 7360
| | | | |
QY 7021 GGAGATGCTCTTCAACAGCTGAGCATTTAAGTCTTTTGCCAGGCCCCCAAGCGGCGAT 7080
| | | | |
Db 7361 GGAGATGCTCTTCAACAGCTGAGCATTTAAGTCTTTTGCCAGGCCCCCAAGCGGCGAT 7420
| | | | |
QY 7081 TCAAGGCTTTTCAACGCGGGGCGGGGCTGCCGATTCGCGAGTCAAGACGCTCTGATGAG 7140
| | | | |
Db 7421 TCAAGGCTTTTCAACGCGGGGCGGGGCTGCCGATTCGCGAGTCAAGACGCTCTGATGAG 7480
| | | | |
QY 7141 TTGGCCCTTTCCGAGACAGGTTTCATCTCTTCAATGCCGCCCTCGAGGGGAGCTTGA 7200
| | | | |
Db 7481 TTGGCCCTTTCCGAGACAGGTTTCATCTCTTCAATGCCGCCCTCGAGGGGAGCTTGA 7540
| | | | |
QY 7201 GATTCAGACCTGAGGCTGAGCAGTGAAGCCCCAACCCGCCCAAGGGGGGAGTGA 7260
| | | | |
Db 7541 GATTCAGACCTGAGGCTGAGCAGTGAAGCCCCAACCCGCCCAAGGGGGGAGTGA 7600
| | | | |
QY 7261 GCTCCGAGCTCGGACCTCGGGGCTCTGCTGTATCTTGCCTGAGAGACCACTCCGCTGTG 7320
| | | | |
Db 7601 GCTCCGAGCTCGGACCTCGGGGCTCTGCTGTATCTTGCCTGAGAGACCACTCCGCTGTG 7660
| | | | |
QY 7321 TGCTGCTCAATGCTATCTTGGAGCGGGGCTTAATACTCTTGTATGCTCCGAAGAG 7380
| | | | |
Db 7661 TGCTGCTCAATGCTATCTTGGAGCGGGGCTTAATACTCTTGTATGCTCCGAAGAG 7720
| | | | |
QY 7381 GAGAGTTACCGATTAACCTCTTGAACACTCCCTGTGCGATATCAACAAAGGTGTAC 7440
| | | | |

Db 7721 GAGAGTTACCGATTAAACCCCTTGAGCACTCCCTGTGGATATGACAACAAGGTGAC 7780
Qy 7441 TGTACCAACAACAAGAGCGGCTCACTAAGGGCTAAAAAGGTAACTTTTGTAGATGACAA 7500
Db 7781 TGTACCAACAACAAGAGCGGCTCACTAAGGGCTAAAAAGGTAACTTTTGTAGATGACAA 7840
Qy 7501 GTGCTGACTCTCTACTACGACTAGTCTTAAAGAACATTAAAGTACGGGCTTCAAGTTC 7560
Db 7841 GTGCTGACTCTCTACTACGACTAGTCTTAAAGAACATTAAAGTACGGGCTTCAAGTTC 7900
Qy 7561 ACCGGAAGGCTCTCAACATGAGAGAGGCTTGCAGTTAAACCCCAACCCCACTTTCGACAA 7620
Db 7901 ACCGGAAGGCTCTCAACATGAGAGAGGCTTGCAGTTAAACCCCAACCCCACTTTCGACAA 7960
Qy 7621 TCTAAATATAGGTTTGGGGCTTAAAGAGGTCGAGGCTTTCGCGAGGCGCTTAAACAC 7680
Db 7961 TCTAAATATAGGTTTGGGGCTTAAAGAGGTCGAGGCTTTCGCGAGGCGCTTAAACAC 8020
Qy 7681 ATCAAGTCGCTGTGGAAGACCTTCGTGAGAGACTCAGAAAACACCAATTCCACACCATTT 7740
Db 8021 ATCAAGTCGCTGTGGAAGACCTTCGTGAGAGACTCAGAAAACACCAATTCCACACCATTT 8080
Qy 7741 ATGCGCAAAAATAGAGGTTCTGCTGAGACCCCAACGAAGGGGCAAGAAAGAGAGTCCG 7800
Db 8081 ATGCGCAAAAATAGAGGTTCTGCTGAGACCCCAACGAAGGGGCAAGAAAGAGAGTCCG 8140
Qy 7801 CTATGCTTAACTCTGACCTCGGCGTCAAGGCTTGCAGAGAGATGAGCCCTTATGACATT 7860
Db 8141 CTATGCTTAACTCTGACCTCGGCGTCAAGGCTTGCAGAGAGATGAGCCCTTATGACATT 8200
Qy 7861 ACACAAAAATTCTCTCAAGCGGTGATGAGGGGCTTTCTTATGATTCAGATATTCGCCGCT 7920
Db 8201 ACACAAAAATTCTCTCAAGCGGTGATGAGGGGCTTTCTTATGATTCAGATATTCGCCGCT 8260
Qy 7921 CAGCGGGTGAAGTTCTCTTGAAGAAGATGGCGGAAAGAAAGAACCTTATGGTTTTTTCG 7980
Db 8261 CAGCGGGTGAAGTTCTCTTGAAGAAGATGGCGGAAAGAAAGAACCTTATGGTTTTTTCG 8320
Qy 7981 TATGATACCCGATGCTTGTGACTCAACCGTCACTGAGAGAGACATCAGACTGAGAGTCC 8040
Db 8321 TATGATACCCGATGCTTGTGACTCAACCGTCACTGAGAGAGACATCAGACTGAGAGTCC 8380
Qy 8041 ATATATCGGGCTGTCTCTTGGCCGAGAGAGGCCCAACTGCACTGCTGACTGAG 8100
Db 8381 ATATATCGGGCTGTCTCTTGGCCGAGAGAGGCCCAACTGCACTGCTGACTGAG 8440
Qy 8101 AGACCTTATGAGGAGGCTTATGTTCAACAGAGAGGCCCAACTGCGGGTACAGGCT 8160
Db 8441 AGACCTTATGAGGAGGCTTATGTTCAACAGAGAGGCCCAACTGCGGGTACAGGCT 8500
Qy 8161 TGCAGGCGCAGCGGGGTGCTCAACCACTAGCATGGGGAACACATCACTACTGTA 8220
Db 8501 TGCAGGCGCAGCGGGGTGCTCAACCACTAGCATGGGGAACACATCACTACTGTA 8560
Qy 8221 GCGTTAGCGGCTTGTAAAGCTGCAAGGATATATGCGGCCCAATGCTGTATGCGGCGAT 8280
Db 8561 GCGTTAGCGGCTTGTAAAGCTGCAAGGATATATGCGGCCCAATGCTGTATGCGGCGAT 8620
Qy 8281 GACTTGTGTCACTCAGAAACCCAGAGGACCGAGAGAGACGAGGGGAACCTGAGAGCC 8340
Db 8621 GACTTGTGTCACTCAGAAACCCAGAGGACCGAGAGAGACGAGGGGAACCTGAGAGCC 8680
Qy 8341 TTACAGGAGGCTATGACCAAGTATTCGCTCTGTGATACCCCAAGACCGAGAT 8400
Db 8681 TTACAGGAGGCTATGACCAAGTATTCGCTCTGTGATACCCCAAGACCGAGAT 8740
Qy 8401 GATCTGAGGCTATTAACATTTGCTCTCAATATGTGTCTGTGCGCTGTGGCCCAAGGCG 8460
Db 8741 GATCTGAGGCTATTAACATTTGCTCTCAATATGTGTCTGTGCGCTGTGGCCCAAGGCG 8800
Qy 8461 CGCGGAGATACATACGAGACCAAGAGCCCAACCACTCCCGCGGGGTGCTGGGAA 8520
Db 8801 CGCGGAGATACATACGAGACCAAGAGCCCAACCACTCCCGCGGGGTGCTGGGAA 8860

Qy 8521 ACAGTTAGACACTCCCTGTCAATTCATGCTGGGAAACATCATCAGTACGCCCGAC 8580
Db 8861 ACAGTTAGACACTCCCTGTCAATTCATGCTGGGAAACATCATCAGTACGCCCGAC 8920
Qy 8581 ATATGGGCTGCGATGCTCTGATGACACACTTTCTCCATTTCTCATGCTCAAGACG 8640
Db 8921 ATATGGGCTGCGATGCTCTGATGACACACTTTCTCCATTTCTCATGCTCAAGACG 8980
Qy 8641 CTGGACCAAGACTCACTTATGATGATGATGAGAGGAGGATGCTCCGTGAGTCCCTTGAC 8700
Db 8981 CTGGACCAAGACTCACTTATGATGATGATGAGAGGAGGATGCTCCGTGAGTCCCTTGAC 9040
Qy 8701 CTCCGACTATATGAAAGGTTAACTGAGGCTTGAAGCTTTTCTCTGACATPACT 8760
Db 9041 CTCCGACTATATGAAAGGTTAACTGAGGCTTGAAGCTTTTCTCTGACATPACT 9100
Qy 8761 CCCCAAGAACTGACACGAGGCTTCAAGCTTCAAGAAACCTTGGGGCGCACCTCTGAC 8820
Db 9101 CCCCAAGAACTGACACGAGGCTTCAAGCTTCAAGAAACCTTGGGGCGCACCTCTGAC 9160
Qy 8821 GCGTGGAGAGCGCGGCACTGACAGTCAAGGAGTCCCTCATCTCCGTTGGGGAGAGG 8880
Db 9161 GCGTGGAGAGCGCGGCACTGACAGTCAAGGAGTCCCTCATCTCCGTTGGGGAGAGG 9220
Qy 8881 GCGTGGAGAGCGCGGCACTGACAGTCAAGGAGTCCCTCATCTCCGTTGGGGAGAGG 8940
Db 9221 GCGTGGAGAGCGCGGCACTGACAGTCAAGGAGTCCCTCATCTCCGTTGGGGAGAGG 9280
Qy 8941 TTGCGGAAAGCGGCTCTCTGATTTATCACTGCTTACCTGTGGGCGCGCGGGGCG 9000
Db 9281 TTGCGGAAAGCGGCTCTCTGATTTATCACTGCTTACCTGTGGGCGCGCGGGGCG 9340
Qy 9001 GACATTTATCAAGAGTGTGCGGTGCGGAGCCCGCTTATGCTCTTGGCTACTCTTA 9060
Db 9341 GACATTTATCAAGAGTGTGCGGTGCGGAGCCCGCTTATGCTCTTGGCTACTCTTA 9400
Qy 9061 CTTTTTGTAGGGGTAGGCTTTTCTTCTACTCCCGCTCGG 9099
Db 9401 CTTTTTGTAGGGGTAGGCTTTTCTTCTACTCCCGCTCGG 9439

RESULT 4
AF177036 9711 bp RNA linear VRL 05-OCT-1999
LOCUS AF177036
DEFINITION Hepatitis C virus clone pJ6CF, complete genome.
ACCESSION AF177036
VERSION AF177036.1 GI:6010579
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE 1 (bases 1 to 9711)
AUTHORS Yanagi, M., Purcell, R.H., Emerson, S.U. and Bukh, J.
TITLE Hepatitis C virus: an infectious molecular clone of a second major genotype (2a) and lack of viability of intertypic 1a and 2a chimeras
JOURNAL Virology 262 (1), 250-263 (1999) Sept. 15, 1999
MEDLINE 99420396
PUBMED 10489358
REFERENCE 2 (bases 1 to 9711)
AUTHORS Bukh, J.
TITLE Direct Subinfection
JOURNAL Submitted (11-AUG-1999) Hepatitis Viruses Section, Laboratory of Infectious Diseases, National Institutes of Allergy and Infectious Diseases, National Institutes of Health, Building 7, Room 201, 7-Center Dr., Bethesda, MD 20892-0740, USA
FEATURES
source
1..9711
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="HC-J6CH"

CDS

```
/db_xref="taxon:11103"
/clone="pJ6CP"
/note="infectious cDNA
genotype: 2a"
341..9442
/codon_start=1
/product="polyprotein"
/db_xref="GI:6010580"
/translation="MSTNRPQRKTKNTNRPODYKPGGGIIVGVVLLPRGPRLL
GVATRTKTSERSQPRGRNPIPKDREKSTGKGPMPVLYGEGIGAMGMLSPG
SRPBGENDRHRSSRNKAVIDLTGCFALMGYIPVAGPLGGVLAALAHVYED
GVNATGNLPGCSFSIFLALLSCITTPAAEAKRNTSTGYMTDNTDNTSIWLOA
AVLHVGCVACEKVGASOCMIYSPVAVNORPGLTGGRTIIMDMKMMSPRTATIL
VGLDGGVMLAOMFIVSPDHMFVQDCKSYEGTITGRMAIDMMKMSPTATML
AVMARPEVILIDISGAHGMFGLAFSQAQMAKAVVILLALAGDAATHVVGSA
AQTGRTLSFDMPGRKIQLVNTGSMHNRNALCNDLHFGTASLSTYKFSNS
GCEBMSACRSIEAFRGMGLQYEDVTPEDMRPYCMHPRQCVASAKTVCGFV
YCFSPVUVVGTDRLAGPYTWGNETDVFLLNSTRPPLGSMFGCTMMMSGVTKC
GAPCRTRADPNASTDLCPTDCPRKHPTTYLKCQSGPMLTPRCLIDYPRLLMHV
TNVYTFKIRMYVGVGRHLTAACNPTGRGNCUEDRQSLSPLHSTTEMLPES
YSLPALSTGLHLHONIVDQPMYGSPLATYIVMEWVILLFLAARCAACLM
MLILQAEALBEKVLIAHAASACNGFLYPIFVAAMVIGKRVPLATYSTGLM
SFLILALPOQAVADSVHGOIGALLVMTLETLTPGYKTLISFLMMLCYLTL
GEAMVQEMAPPMOVRGGDGIWAVAFYGVFEDITKMLAVGPAVYLKGLATRP
YFVRAHALBMCTMAHLAGRYVOMALLGRWGYIVDHLTPMSDMAKGLRDLA
VAVPPIIFSPMEKVIYMGAEYACDGIHLGIPVSAHGEVILGADGTTSKMSLL
APITAAQOTRGLGTLIVSMGRDKTEQAGEIQLVSTVRSFGLSTISGLMTVHG
AGNKTLAGSRSPVOTMYSAGDLVGMSPGPKSLPCTCGAVDLVLTNRMDVIPA
RRGDKRGLASPRPLSTLKGSGGPLYCRGHAVGFRVAVCSRGAKSIDIPVET
LDIVTRSPFSDNSTPAPVOTYOVGYLHAPGSGSKTKVVAAGYKVLVNPV
AATLGRGAVLSKAGINPNIRGTGRTVTGAPIYTSYGFELADGCGAGAYDIIICD
ECHANSTTILIGITVLDQAEYKGLVLTATATPRESVTHPHNIEVALGQGEIP
FYRAPIPLSTIKGRHLIFCHSKKCCDELAALRGMNLVAAYRGDVAVIPQGGV
VVVATDALMTGYGDPDSVIDCNVAVTQVDFSLDPTFTITQIIVQDAVRSQRRR
TGRGLGIIYRVSTGERASGMPDSVLCEDAGAAVELTPSETVRLRAYFNTPL
PCODLHEFWEAVPTGLTHIDAPHLSTQSGENPALTVOATVCAKAPAPMPV
MCKLRKRPKTLVGPRLYRLGSLVNRHLYTHPVYIATCMQADLEWNTWVLG
GVLAAVATCLATGCVCTIGRLHINORAVAPREVEYEAPEDEKASDALIEBQ
RIAEMLKSKIOGLLOQASQKQADIQTPVAVSPREVEYEAPEDEKASDALIEBQ
LSPNPAVSMMAASALTSPSTSTILLNLGCMWLASQEPAGATFVAGSIVVGA
VGSIGKVLVLDILAGGAGISGALVAFKMSGEKPEMEVNLPLPILSGALVAV
ICAAILRHVGPGEAGVOMNRLIAFASRNHYAPTHYVESASQVOMDLSLTT
SLRLRLHMTDEDCPIPCGSMRLDVAWVCTIILDKMLTKSLPRKMGKLPVSCQ
KGYGVAGTGIMTTRPCGANISGNRLSMKRTGKTKMNIQGFPIINCTBEGC
VPRPAPNPKFALIRVAASEAVTQHSYHITGLTNDNLKVCQSPSPFVNDVAY
QIRFAPTPKPFREDVSCVGLNSFVVSQLPDDEPDVMSMLTDSHITAEFA
ARLRAGSPSEASASQSLASPLRATCTTHGKAYVDVMDANLPMGCVTRTIESG
KVVLVLDPMVERSDLEPSIPSEVMLPKRFPALPAARADYNPLVESKRPDY
QPATVAGCALPPRKTPTPPRRRRTVGSSEDSIGALQOLAKSRQPPSPDSGSL
TGGAADSGSQTPPELALSETGISISMPLEBGLDPELPROVERQPPQGVAVP
GSDGSWSTCSBEDDVSVCSSMSYSWGLDITPCSPEEKLPIINPLSNLRHANYK
CTTTKSASLAKKVTPEPMQVLDYSYDVKDILKASKTATALLTMEBACQTPPS
ARXCYGKAKEVRSLSGRAVNIHKSWKDLDESEPTIPTTIAMKNEFVCDPTPKGS
KARLIYVPLIGVRCERKMLYDITOKLPQAVGASGFQYSPORAREFLIYGMENK
DMPGSDYDRCPDSTVTERDITPESIVYRCSLPERAHYIHSLTRLYXGMENK
GQTCGRRCASGYLTTSMGNITTCYKALACCAAGIAPMLVCGDDIVISSESG
TEEDERNLRAFTAMTRYSAPEQDPPPEYDEBLITSCSNVAFESILIMADTDONIN
FEMYGAVYSPLDPAIIERLHGLDAFSLHATYPHETLVAALRGLAPRAMKS
BARAVNASTLSRGRGAVCGRYLFNMAVVKTKLTPLEARLLDLSWFLVAGCGDI
YHSVSRAPRLILFGLLLFVGVGLPLPAR"
```

ORIGIN

```
Query Match 100.0%; Score 9099; DB 14; Length 9711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9099; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGCAAAATCTTAACCTCAAGAAAAACCAAGAACCAACCGTCCGCCACAA 60
Db 341 ATGAGCAAAATCTTAACCTCAAGAAAAACCAAGAACCAACCGTCCGCCACAA 400
QY 61 GAGCTTAAGTTTCGGGCGGCGGCGAGATGTTGGCGAGATATCTTGTCGCGGAG 120
```

```
|||||
Db 401 GAGCTTAAGTTTCGGGCGGCGGCGGCGAGATGTTGGCGAGATATCTTGTCGCGGAG 460
QY 121 GGGCCCAAGTTCGGGTGTGCGCGGACCAAGGAAGACTTCGAGCGGTCCAGCACTGGA 180
Db 461 GGGCCCAAGTTCGGGTGTGCGCGGACCAAGGAAGACTTCGAGCGGTCCAGCACTGGA 520
QY 181 AGGCGCCAGCCCATTCCTTAAGATCGGCGCTCCACTGGAACAAATCTTGGGGAAAAACGGA 240
Db 521 AGGCGCCAGCCCATTCCTTAAGATCGGCGCTCCACTGGAACAAATCTTGGGGAAAAACGGA 580
QY 241 TACCCCTGAGCCCTATACGGGAATGAGGAGCTCGGCTGAGGACAGATGGCTCTGCTCCCC 300
Db 581 TACCCCTGAGCCCTATACGGGAATGAGGAGCTCGGCTGAGGACAGATGGCTCTGCTCCCC 640
QY 301 CGAGTTCCTCGCTCTTGGGAGCCCAATGACCCCGGATAGTTCGCGAACGTGGGT 360
Db 641 CGAGTTCCTCGCTCTTGGGAGCCCAATGACCCCGGATAGTTCGCGAACGTGGGT 700
QY 361 AAGTCATCGATACCTTAACGTGCGGCTTGGCGACCTCATGGGGTACATCCCTGTCGTG 420
Db 701 AAGTCATCGATACCTTAACGTGCGGCTTGGCGACCTCATGGGGTACATCCCTGTCGTG 760
QY 421 GGGCGCCCGCTCGGCGGCGTGGCAGAGCTCTGCGCATGGCGTGAAGTCTGAGAGAC 480
Db 761 GGGCGCCCGCTCGGCGGCGTGGCAGAGCTCTGCGCATGGCGTGAAGTCTGAGAGAC 820
QY 481 GGGGTAAATTTTGAACAGGAACTTACCCGTTGCTCTTTTCTATCTTCTTGGTGGC 540
Db 821 GGGGTAAATTTTGAACAGGAACTTACCCGTTGCTCTTTTCTATCTTCTTGGTGGC 880
QY 541 CTGCTGCTCGCATACACCCCGGTCTCCGCTCGGAAGTGAAGAACATCAGTACCGGC 600
Db 881 CTGCTGCTCGCATACACCCCGGTCTCCGCTCGGAAGTGAAGAACATCAGTACCGGC 940
QY 601 TACATGTCATGACTACGACTGCAACCAATGACAGATTACTGGAGCTCCAGGCTGCTGC 660
Db 941 TACATGTCATGACTACGACTGCAACCAATGACAGATTACTGGAGCTCCAGGCTGCTGC 1000
QY 661 CTCACGTCCTCCGGGTGCGTCCCGTGGAGAAATGGGGAATGATCTCACTGCTGGATA 720
Db 1001 CTCACGTCCTCCGGGTGCGTCCCGTGGAGAAATGGGGAATGATCTCACTGCTGGATA 1060
QY 721 CCGGTCTACCGCAATGGCGGTGACGGGCGCGGCGGCGCTTACGAGCGGCTTGGCGGAC 780
Db 1061 CCGGTCTACCGCAATGGCGGTGACGGGCGCGGCGGCGCTTACGAGCGGCTTGGCGGAC 1120
QY 781 CACATGCAATGATGATGATGTCGCGACAGCTGCTGCTGCTCACTGAGGGGACCTC 840
Db 1121 CACATGCAATGATGATGATGTCGCGACAGCTGCTGCTGCTCACTGAGGGGACCTC 1180
QY 841 TGGGTGGGGTGAATGCTCGGAGCCCAATGTTGATGTTCTGCGCGACACACTGGTTT 900
Db 1181 TGGGTGGGGTGAATGCTCGGAGCCCAATGTTGATGTTCTGCGCGACACACTGGTTT 1240
QY 901 GTCCAGACTGCAATGCTGCTCACTACCTCGTGAACATCATGGAACACCGCATGGCATGG 960
Db 1241 GTCCAGACTGCAATGCTGCTCACTACCTCGTGAACATCATGGAACACCGCATGGCATGG 1300
QY 961 GACATGATGATGATGATGTCGCGACAGCTACCATGATCTTGGCGTACGCGATGCGTGTG 1020
Db 1301 GACATGATGATGATGATGTCGCGACAGCTACCATGATCTTGGCGTACGCGATGCGTGTG 1360
QY 1021 CCGGAGGTCATTATGACATCATTTAGCGGGGCTCATTTGGGGGCTCATTTGGGGGCTTGGCC 1080
Db 1361 CCGGAGGTCATTATGACATCATTTAGCGGGGCTCATTTGGGGGCTCATTTGGGGGCTTGGCC 1420
QY 1081 TACTTCTCTATGAGGAGGCTGGGGAAGTGTGTGATCTCTTCTGTTGGCGCGGAG 1140
Db 1421 TACTTCTCTATGAGGAGGCTGGGGAAGTGTGTGATCTCTTCTGTTGGCGCGGAG 1480
QY 1141 GTGAGCGCGGCAACCATGCTGTTGGGGGTTCTGCGCGGACCAACCGGGCGCTTACC 1200
```

Db 1481 GTGACGGCGACCCCACTACTGTGTGGGGTTCTGCCGGCAGAACACCGGGCGCTCAC 1540
Oy 1201 AGCTTATTTGACATGAGGCCCCAGAGAGAAAATCCAGCTCGTTAACACCAATAGGAGCTGG 1260
Db 1541 AGCTTATTTGACATGAGGCCCCAGAGAGAAAATCCAGCTCGTTAACCAATAGGAGCTGG 1600
Oy 1261 CACATCAACCGCACCGCCCTGAACTGCATGACTCTTTGCACACCGGCTTTATCGGCTCT 1320
Db 1601 CACATCAACCGCACCGCCCTGAACTGCATGACTCTTTGCACACCGGCTTTATCGGCTCT 1660
Oy 1321 CTGTTCTTACACCAAGCTTCAACTCGTCAGAGATGTCCCGAACCATATGTCGGCTGCGC 1380
Db 1661 CTGTTCTTACACCAAGCTTCAACTCGTCAGAGATGTCCCGAACCATATGTCGGCTGCGC 1720
Oy 1381 AGTATCGAGGCTTCCGGGTGGAGATGGGGCGCTTGGCAATATGAGGATTAATGTCCCAAT 1440
Db 1721 AGTATCGAGGCTTCCGGGTGGAGATGGGGCGCTTGGCAATATGAGGATTAATGTCCCAAT 1780
Oy 1441 CCAGAGGATATGAGACCTTATTTGCTGGCACTAACCCAGAGGAGTGTGGCTGTCTCC 1500
Db 1781 CCAGAGGATATGAGACCTTATTTGCTGGCACTAACCCAGAGGAGTGTGGCTGTCTCC 1840
Oy 1501 GCGAAGACTGTGTGGGCGCACTGTACTGTTCACCCCGAGCCAGTGTGTGGGACG 1560
Db 1841 GCGAAGACTGTGTGGGCGCACTGTACTGTTCACCCCGAGCCAGTGTGTGGGACG 1900
Oy 1561 ACCGACAGGCTTGGAGCGCCCACTTACAGTGGGGGAGAAATGAGACAGATGTCTCTA 1620
Db 1901 ACCGACAGGCTTGGAGCGCCCACTTACAGTGGGGGAGAAATGAGACAGATGTCTCTA 1960
Oy 1621 TTGAAACAGACTCGACACCGCTGGGGTCAATGTTCGGCTGCACGTGTGAATGAATCTTCT 1680
Db 1961 TTGAAACAGACTCGACACCGCTGGGGTCAATGTTCGGCTGCACGTGTGAATGAATCTTCT 2020
Oy 1681 GGCTACACCAAGACTTGGGGGCAACACCTGGCGTACTAGAGCTTCAAGCGGACG 1740
Db 2021 GGCTACACCAAGACTTGGGGGCAACACCTGGCGTACTAGAGCTTCAAGCGGACG 2080
Oy 1741 ACGACACTGTGTGGCCCAACGGAAGCTGTTTATGAGAGCATCTGTATACCACTTACCTCAA 1800
Db 2081 ACGACACTGTGTGGCCCAACGGAAGCTGTTTATGAGAGCATCTGTATACCACTTACCTCAA 2140
Oy 1801 TGGGCTCTGGGCTCTGGCTCAACGCAAGTGCCTGATCGACTACCCCTACAGGCTCTGG 1860
Db 2141 TGGGCTCTGGGCTCTGGCTCAACGCAAGTGCCTGATCGACTACCCCTACAGGCTCTGG 2200
Oy 1861 CATTACCCCTGCACAGTTAATCATATACACTTTCAAAATTAAGATGTATGTGGAGGGGTT 1920
Db 2201 CATTACCCCTGCACAGTTAATCATATACACTTTCAAAATTAAGATGTATGTGGAGGGGTT 2260
Oy 1921 GAGCAGAGCTCAACGGCTGCATGCAATTTCACTGGTGGGATCGTTCGAATCTTGGAGAC 1980
Db 2261 GAGCAGAGCTCAACGGCTGCATGCAATTTCACTGGTGGGATCGTTCGAATCTTGGAGAC 2320
Oy 1981 AGAGCAGAAATGCACTGTCTCTTGTGTGCACTTCACAGGAAATGGGCCATTTTACT 2040
Db 2321 AGAGCAGAAATGCACTGTCTCTTGTGTGCACTTCACAGGAAATGGGCCATTTTACT 2380
Oy 2041 TGCTCTTACTCGGACCTGCGCGCTGTGCACTGGTCTTCTCAACCTGCACAAACATC 2100
Db 2381 TGCTCTTACTCGGACCTGCGCGCTGTGCACTGGTCTTCTCAACCTGCACAAACATC 2440
Oy 2101 GTGACGTAACAATTCATGTATGTGCCTATCACCTGCCCTCACAAAATACATGTGCGAATGG 2160
Db 2441 GTGACGTAACAATTCATGTATGTGCCTATCACCTGCCCTCACAAAATACATGTGCGAATGG 2500
Oy 2161 GAGTGGGTAATTAATCTTATATCTGTCTTAAAGCGACGCAAGGTTTGGGCTGCTTATGG 2220
Db 2501 GAGTGGGTAATTAATCTTATATCTGTCTTAAAGCGACGCAAGGTTTGGGCTGCTTATGG 2560
Oy 2221 ATGTCACTTGTGGGCGAGGCGAGAGAGAGACTAGAGAGAGTGTGATCTTGGACGCT 2280
Db 2561 ATGTCACTTGTGGGCGAGGCGAGAGAGAGACTAGAGAGAGTGTGATCTTGGACGCT 2620

Oy 2281 GCGAGCGAGTATAGTCAATGAGCTTCTATATTTTGTATCATCTTTTCTGGCTGCTTGG 2340
Db 2621 GCGAGCGAGTATAGTCAATGAGCTTCTATATTTTGTATCATCTTTTCTGGCTGCTTGG 2680
Oy 2341 TACATCAAGGTCGGGTAGTCCCTTAACTTACCTTATCCCTCACTGGCTGTGTCTTT 2400
Db 2681 TACATCAAGGTCGGGTAGTCCCTTAACTTACCTTATCCCTCACTGGCTGTGTCTTT 2740
Oy 2401 AGCTTCTGCTCTTACATTTGCCCAAGAGCTTATGCTTATGATGACATCTGTGCATGGC 2460
Db 2741 AGCTTCTGCTCTTACATTTGCCCAAGAGCTTATGCTTATGATGACATCTGTGCATGGC 2800
Oy 2461 CAGATGAGAGCGGCTGTGCTGTATGATCACTCTTTTACTCTACCCCGAGTAAAG 2520
Db 2801 CAGATGAGAGCGGCTGTGCTGTATGATCACTCTTTACTCTACCCCGAGTAAAG 2860
Oy 2521 ACCCTTCTACGCGGTTTTTGTGTGGTGTGTATCTTCTGACCTGTGGGGAAAGCTAG 2580
Db 2861 ACCCTTCTACGCGGTTTTTGTGTGGTGTGTATCTTCTGACCTGTGGGGAAAGCTAG 2920
Oy 2581 GTCAGAGTGGGCAACCACTATGACAGTGGCGGCTGTGATGACATTAATGAGGCT 2640
Db 2921 GTCAGAGTGGGCAACCACTATGACAGTGGCGGCTGTGATGACATTAATGAGGCT 2980
Oy 2641 GTGCGCAATATTTCAACCAAGTGTGTGTGACATTAACCAAGTGTCTTGGCGGCTT 2700
Db 2981 GTGCGCAATATTTCAACCAAGTGTGTGTGACATTAACCAAGTGTCTTGGCGGCTT 3040
Oy 2701 GGGCTCTGCTTACCTCTTAAAGTGTCTTGAAGCGGCTGCTTCTGTCAAGGCTCAC 2760
Db 3041 GGGCTCTGCTTACCTCTTAAAGTGTCTTGAAGCGGCTGCTTCTGTCAAGGCTCAC 3100
Oy 2761 GCTCTACTAGAGATGTGACCAATGCAAGGACCTTGCAGGGGGAGTACGTCCAGATG 2820
Db 3101 GCTCTACTAGAGATGTGACCAATGCAAGGACCTTGCAGGGGGAGTACGTCCAGATG 3160
Oy 2821 GCGCTACTAGCCTTGGAGTGGACTGTGCACTTATGATGACCACTTACCCCTATG 2880
Db 3161 GCGCTACTAGCCTTGGAGTGGACTGTGCACTTATGATGACCACTTACCCCTATG 3220
Oy 2881 TCGGATTTGGGCTGTAGTGGCTTGGGGACCTGGCGGTGGCGGTATGACTTATCTTC 2940
Db 3221 TCGGATTTGGGCTGTAGTGGCTTGGGGACCTGGCGGTGGCGGTATGACTTATCTTC 3280
Oy 2941 AGTCGATGGAAGAAAGTCAATTTGTCTGGGGAGCGGAGACAGCTCTTGTGGGACATT 3000
Db 3281 AGTCGATGGAAGAAAGTCAATTTGTCTGGGGAGCGGAGACAGCTCTTGTGGGACATT 3340
Oy 3001 TTACAGGACCTTCCGCTGTCCGCGCACTTGTGTGGAGGTCTCTTGGCCGACTGAT 3060
Db 3341 TTACAGGACCTTCCGCTGTCCGCGCACTTGTGTGGAGGTCTCTTGGCCGACTGAT 3400
Oy 3061 GGGCTATACCTCCAAAGGGGTGGAGTCTTCCGGCCCATCATGCTTATGCGCCAGAGACA 3120
Db 3401 GGGCTATACCTCCAAAGGGGTGGAGTCTTCCGGCCCATCATGCTTATGCGCCAGAGACA 3460
Oy 3121 CGTGGCTTTTGGGCAACATAGTGTAGACATGACGGGCGGCAACAAGAGAGCT 3180
Db 3461 CGTGGCTTTTGGGCAACATAGTGTAGACATGACGGGCGGCAACAAGAGAGCT 3520
Oy 3181 GGGGAAATTCAGGTCCTGTTCACAGTCACTCACTCTTCTTGGAAATTCATCTGGGG 3240
Db 3521 GGGGAAATTCAGGTCCTGTTCACAGTCACTCACTCTTCTTGGAAATTCATCTGGGG 3580
Oy 3241 GTTTTGTGACCTGTCTACATGAGTGTGAGCAACAACCTTGGCGGGCTTCAAGGGGTCCG 3300
Db 3581 GTTTTGTGACCTGTCTACATGAGTGTGAGCAACAACCTTGGCGGGCTTCAAGGGGTCCG 3640
Oy 3301 GTACCGCAGATGTACTCAAGTGTGAGGGGAACTTATGAGGGTGGCCAGCCCCCTGGG 3360
Db 3641 GTACCGCAGATGTACTCAAGTGTGAGGGGAACTTATGAGGGTGGCCAGCCCCCTGGG 3700

QY 3361 ACTAAATCTTTGAGACCGGTGACGCTGTGAGCGGTGACCTGTACTGTGACCGCGAAG 3420
DB 3701 ACTAAATCTTTGAGACCGGTGACGCTGTGAGCGGTGACCTGTACTGTGACCGCGAAG 3760
QY 3421 GCTGATGTCAATCCCGGCTCGAAGACCGGGGACAAACCGGGAGCGCTACTCTCCCGAGA 3480
DB 3761 GCTGATGTCAATCCCGGCTCGAAGACCGGGGACAAACCGGGAGCGCTACTCTCCCGAGA 3820
QY 3481 CCTCTTTCACCTTGAAGGGGTCTCGAAGAGCGCGGTGCTATGCCCCAGGGGCAAGCT 3540
DB 3821 CCTCTTTCACCTTGAAGGGGTCTCGAAGAGCGCGGTGCTATGCCCCAGGGGCAAGCT 3880
QY 3541 GTGAGACTCTTCCGGGAGCGCTGTGTCTCTCGGGCGGTGTAGTCAATGATTTGATC 3600
DB 3881 GTGAGACTCTTCCGGGAGCGGTGTGTCTCTCGGGCGGTGTAGTCAATGATTTGATC 3940
QY 3601 CCGGTTGAGACACTGACATGTGTACCGCGGTCCCGACCTTTAGTGAACAAGACACCA 3660
DB 3941 CCGGTTGAGACACTGACATGTGTACCGCGGTCCCGACCTTTAGTGAACAAGACACCA 4000
QY 3661 CCGTGTGTGCCCCGACCTATAGGTGGGTACTGTGATGCCCGACCTGGCAGTGGGAAG 3720
DB 4001 CCGTGTGTGCCCCGACCTATAGGTGGGTACTGTGATGCCCGACCTGGCAGTGGGAAG 4060
QY 3721 AGCACCAAGTTCCTGTGCGATATGCTCTCAGGGGTATAAAGTGTAGTCTTAATCCC 3780
DB 4061 AGCACCAAGTTCCTGTGCGATATGCTCTCAGGGGTATAAAGTGTAGTCTTAATCCC 4120
QY 3781 TCAGTGCTGCCACCTGTGGGTTTGGGGCGTACTTGTCTAAAGGACATGGCATATCCC 3840
DB 4121 TCAGTGCTGCCACCTGTGGGTTTGGGGCGTACTTGTCTAAAGGACATGGCATATCCC 4180
QY 3841 AACATTAGGACTGAGTCAAGGACTGTGACGACCGGGGCGCCCATCAAGTCTCCACATAT 3900
DB 4181 AACATTAGGACTGAGTCAAGGACTGTGACGACCGGGGCGCCCATCAAGTCTCCACATAT 4240
QY 3901 GGGAAATTCCTGCGCAGTGGGGGCTGTGCGGGCGGCGCTTACGACATCATATGTGAT 3960
DB 4241 GGGAAATTCCTGCGCAGTGGGGGCTGTGCGGGCGGCGCTTACGACATCATATGTGAT 4300
QY 3961 GAAATGCCATCCGTGACCTTACCAACCATCTTGGCATGGAAACAGTCTTTGATCAAGCA 4020
DB 4301 GAAATGCCATCCGTGACCTTACCAACCATCTTGGCATGGAAACAGTCTTTGATCAAGCA 4360
QY 4021 GAAGACACTGGGGTCAACCTAACTGTGTGCTGACGCTACGCCCCCTGGGTGAGTACACA 4080
DB 4361 GAAGACACTGGGGTCAACCTAACTGTGTGCTGACGCTACGCCCCCTGGGTGAGTACACA 4420
QY 4081 ACCCCCCACCCCAACATAGAGAGAGTGGCCCTTGGGAGAGAGGCGAGATCCCCTTCTAT 4140
DB 4421 ACCCCCCACCCCAACATAGAGAGAGTGGCCCTTGGGAGAGAGGCGAGATCCCCTTCTAT 4480
QY 4141 GGGAGGGGCAATTCCTGTCTTACATCAAGGAGAAAGACATCTGATCTTCTGCAATTC 4200
DB 4481 GGGAGGGGCAATTCCTGTCTTACATCAAGGAGAAAGACATCTGATCTTCTGCAATTC 4540
QY 4201 AAAAAAAGTGTACGAGCTCGGGGCGGCGCTTGGGGTATGGGCTTGAATCTAGTGGCA 4260
DB 4541 AAAAAAAGTGTACGAGCTCGGGGCGGCGCTTGGGGTATGGGCTTGAATCTAGTGGCA 4600
QY 4261 TACTACAGAGGGTGTGACGCTCTCCGTAAATCAAACTCAGGGAGAGCTAGTGTGTGCGC 4320
DB 4601 TACTACAGAGGGTGTGACGCTCTCCGTAAATCAAACTCAGGGAGAGCTAGTGTGTGCGC 4660
QY 4321 ACCGACGCTCATGACAGGGTATATCTGGGACCTTTGACTCCGTGATGCACTGCAACGTA 4380
DB 4661 ACCGACGCTCATGACAGGGTATATCTGGGACCTTTGACTCCGTGATGCACTGCAACGTA 4720
QY 4381 GGGGTCACTCAAGTGTAGACTTCAAGTTAGACCCCACTTCAACCATTAACAACAAT 4440
DB 4721 GGGGTCACTCAAGTGTAGACTTCAAGTTAGACCCCACTTCAACCATTAACAACAAT 4780
QY 4441 GTCCCTCAAGACGCTGTCTGACGTAGCCAGCGCGGGGTGCAACGGGTAGGGAGAACTG 4500

DB 4781 GTCCCTCAAGACGCTGTCTGACGTAGCCAGCGCGGGGTGCAACGGGTAGGGAGAACTG 4840
QY 4501 GGCATTTATAGATATGTTTCACTGTGAGCGAGCTCAGAAATGTTTGAACATGTAGT 4560
DB 4841 GGCATTTATAGATATGTTTCACTGTGAGCGAGCTCAGAAATGTTTGAACATGTAGT 4900
QY 4561 CTCTGTAGTGTCAACGACGAGGGGCGGATGGTATGAGTCAACCATGGAGAACACC 4620
DB 4901 CTCTGTAGTGTCAACGAGGGGCGGATGGTATGAGTCAACCATGGAGAACACC 4960
QY 4621 GTGAGGCTCAGGGCGTATTTCAACACGCGCGGTTTGCCTGTGTGCCAAGACCATTTGAG 4680
DB 4961 GTGAGGCTCAGGGCGTATTTCAACACGCGCGGTTTGCCTGTGTGCCAAGACCATTTGAG 5020
QY 4681 TTTTGGAGGACGTTTTCACCGGCTTCACACATATAGTCCCATCTTCTCCCAACA 4740
DB 5021 TTTTGGAGGACGATTTTTCACCGGCTTCACACATATAGTCCCATCTTCTCCCAACA 5080
QY 4741 AAGCAATCGGGGGAAAAATTTGCGATCTTAAACAGCCCTACAGGCTACAGTGGCGTAGG 4800
DB 5081 AAGCAATCGGGGGAAAAATTTGCGATCTTAAACAGCCCTACAGGCTACAGTGGCGTAGG 5140
QY 4801 GCCAAAGCCCCCCCCCGTCTGTGAGCGTATGTGAGTGTGACTCGACTCAAGCCC 4860
DB 5141 GCCAAAGCCCCCCCCCGTCTGTGAGCGTATGTGAGTGTGACTCGACTCAAGCCC 5200
QY 4861 ACACTGTGGGCCCCACACTCTCTGTATACCGCTTGGGCTGTGTAAACAGAGTCAAC 4920
DB 5201 ACACTGTGGGCCCCACACTCTCTGTATACCGCTTGGGCTGTGTAAACAGAGTCAAC 5260
QY 4921 CTCACACATCCGTTGACGAAATATCATCGCCACCTGATGGAAGCCGACCTTTGAGTATG 4980
DB 5261 CTCACACATCCGTTGACGAAATATCATCGCCACCTGATGGAAGCCGACCTTTGAGTATG 5320
QY 4981 ACCAGACATGGGTCTTGTGAGGGGAGTCTTGGCGGCGCTGCGCGGTATTCCTGGCG 5040
DB 5321 ACCAGACATGGGTCTTGTGAGGGGAGTCTTGGCGGCGCTGCGCGGTATTCCTGGCG 5380
QY 5041 ACCGGGTGTGTTCATCATCGGCGCTTGACATTAACCAAGAGCGTGTGGCGCG 5100
DB 5381 ACCGGGTGTGTTCATCATCGGCGCTTGACATTAACCAAGAGCGTGTGTGGCGCG 5440
QY 5101 GACAAGAGGTCTCTATAGAGCTTTGATGAGATGGAAGAGTGTCTTAGGGCGGCT 5160
DB 5441 GACAAGAGGTCTCTATAGAGCTTTGATGAGATGGAAGAGTGTCTTAGGGCGGCT 5500
QY 5161 CTCATTGAAGAGGGGAGCGAGATAGCCGAGATCTGAAGTCCAAGATCCAAGGCTTATTTG 5220
DB 5501 CTCATTGAAGAGGGGAGCGAGATAGCCGAGATCTGAAGTCCAAGATCCAAGGCTTATTTG 5560
QY 5221 CAGCAAGCTTCAACAAGCTCAAGACATTAACAACCACTGTGAGGCTTATGCGCCAG 5280
DB 5561 CAGCAAGCTTCAACAAGCTCAAGACATTAACAACCACTGTGAGGCTTATGCGCCAG 5620
QY 5281 GTAGAACATTCGAGGCGAAACATAGTGAATCTTATTAAGGGGCAATCCAATACCTGCA 5340
DB 5621 GTAGAACATTCGAGGCGAAACATAGTGAATCTTATTAAGGGGCAATCCAATACCTGCA 5680
QY 5341 GGACTATCAACACTGACAGGGAACCTGACGATGCTTCAATGATGAGCGTTCAGTGGCGCC 5400
DB 5681 GGACTATCAACACTGACAGGGAACCTGACGATGCTTCAATGATGAGCGTTCAGTGGCGCC 5740
QY 5401 CTCAACAATCGCTGTCAACAAGACACATATCTTCTCAACAATTTTGGGGGCGCTGCTA 5460
DB 5741 CTCAACAATCGCTGTCAACAAGACACATATCTTCTCAACAATTTTGGGGGCGCTGCTA 5800
QY 5461 GCATCCCAATTTGACACACCGCGGGGCGACGAGCTGTGTGTGCTAGTGGCTAGTGGGA 5520
DB 5801 GCATCCCAATTTGACACACCGCGGGGCGACGAGCTGTGTGTGCTAGTGGCTAGTGGGA 5860
QY 5521 GTGCGGTGAGCAGTATAGCTTAAAGTGTCTAGTGAACAATCTGGCAGGGATAGT 5580

Db 5861 GCTGCGTAGGAGATAGAGCTTAGTAGAGGTGCTAGTAGACATCTTGCGAGGGTATGCT 5920
OY 5881 GCGGGCATTTTGCGGGGGCTCTGCTGSCATTTCAAGATCATGCTGCGAGAAAGCCCTCCATG 5640
Db 5921 GCGGGCATTTTGCGGGGGCTCTGCTGSCATTTCAAGATCATGCTGCGAGAAAGCCCTCCATG 5980
OY 5641 GAGATGTCGTCACTTGCTGCTGCGAAATTCGTCTCCGGGTGCTTGATAGTGGAGTC 5700
Db 5981 GAGATGTCGTCACTTGCTGCTGCGAAATTCGTCTCCGGGTGCTTGATAGTGGAGTC 6040
OY 5701 ATGCGCGGGGCATTTGCGCGCGACACGTGGGACCGGGGGAAAGCGGCTGCATATGATG 5760
Db 6041 ATGCGCGGGGCATTTGCGCGCGACACGTGGGACCGGGGGAAAGCGGCTGCATATGATG 6100
OY 5761 AATAGACTCATGCTTTGCTTCCAGAGAAATCACTGCGCCCCACCACTAGTGCAG 5820
Db 6101 AATAGACTCATGCTTTGCTTCCAGAGAAATCACTGCGCCCCACCACTAGTGCAG 6160
OY 5821 GAGTCGGAATGCGTCGCGAGCGGTGTGACCAACTACTTGGCTCCCTTACCATTAACGAGCTG 5880
Db 6161 GAGTCGGAATGCGTCGCGAGCGGTGTGACCAACTACTTGGCTCCCTTACCATTAACGAGCTG 6220
OY 5881 CTCGAGAAGCTCCAGAACTGATTACTGAGACTGCCCCCATTCGATGGCGGGCTCGTG 5940
Db 6221 CTCGAGAAGCTCCAGAACTGATTACTGAGACTGCCCCCATTCGATGGCGGGCTCGTG 6280
OY 5941 CTCGCGAGATGTCGAGACTGAGGTTTGACCACTCTCAACAGACTTTAAATTTGGCTGACC 6000
Db 6281 CTCGCGAGATGTCGAGACTGAGGTTTGACCACTCTCAACAGACTTTAAATTTGGCTGACC 6340
OY 6001 TCCTAATTTATTCCTCAAGATGCTCGGCTCCCTTTGTCTCTGTCAAAAGGGGTACAG 6060
Db 6341 TCCTAATTTATTCCTCAAGATGCTCGGCTCCCTTTGTCTCTGTCAAAAGGGGTACAG 6400
OY 6061 GCGGTGCGGGCCGGGACTGAGACTCATGACCAAGGTGCTTGCGGGGCCAATATCTCT 6120
Db 6401 GCGGTGCGGGCCGGGACTGAGACTCATGACCAAGGTGCTTGCGGGGCCAATATCTCT 6460
OY 6121 GGCATATGTCGCTTGCGGCTCTCATGAGATCAAGGGGCTTAAGACTGATGATATCTG 6180
Db 6461 GGCATATGTCGCTTGCGGCTCTCATGAGATCAAGGGGCTTAAGACTGATGATATCTG 6520
OY 6181 CAGGGGACCTTTCTCATATGTTTACACGAGGGCCAGTGGTCCGAAACCGCGCCA 6240
Db 6521 CAGGGGACCTTTCTCATATGTTTACACGAGGGCCAGTGGTCCGAAACCGCGCCA 6580
OY 6241 AACTTTAAGGTGCGCATGAGAGGTGCGGCTGAGAGTACGGGGAAGTACAGAGAC 6300
Db 6581 AACTTTAAGGTGCGCATGAGAGGTGCGGCTGAGAGTACGGGGAAGTACAGAGAC 6640
OY 6301 GGGTCATACCACTAGATTAACAGAGCTCACTGATTAATTTGAAGTCCCTGCGCACTA 6360
Db 6641 GGGTCATACCACTAGATTAACAGAGCTCACTGATTAATTTGAAGTCCCTGCGCACTA 6700
OY 6361 CCCTCTCCGAGTTCTTTCTGCGGTGACGAGAGTGCAGATCATAGTGTGGCCCA 6420
Db 6701 CCCTCTCCGAGTTCTTTCTGCGGTGACGAGAGTGCAGATCATAGTGTGGCCCA 6760
OY 6421 CGAAGCGGTTTTCCGGGATGAGAGTCTGCTGCGGCTTAATTCATTTGTCGTC 6480
Db 6761 CGAAGCGGTTTTCCGGGATGAGAGTCTGCTGCGGCTTAATTCATTTGTCGTC 6820
OY 6481 GGGTCACAGCTTCTTTGCGAGCCCTGAAACCGACACAGAGTATGATGCTCATCA 6540
Db 6821 GGGTCACAGCTTCTTTGCGAGCCCTGAAACCGACACAGAGTATGATGCTCATCA 6880
OY 6541 GATTCATCTCATATACCGCGGAGACTGACGCGGCTTTAGCGCGGGGTCAACCCCA 6600
Db 6881 GATTCATCTCATATACCGCGGAGACTGACGCGGCTTTAGCGCGGGGTCAACCCCA 6940
OY 6601 TCCGAGGCAAGCTCCCTCGCGAGCCAGAGTATGAGGACCATGCTGCGAGCCACTGAC 6660
Db 6941 TCCGAGGCAAGCTCCCTCGCGAGCCAGAGTATGAGGACCATGCTGCGAGCCACTGAC 7000

OY 6661 ACCCAAGGCAAGCTTATGATGTGACATGTGTGATGCTTAACCTGTTCAAGGGCGAT 6720
Db 7001 ACCCAAGGCAAGCTTATGATGTGACATGTGTGATGCTTAACCTGTTCAAGGGCGAT 7060
OY 6721 GTGACTCGATAGAGTCTGCGGTCCAAAGTGTGCTTCTGAGACTCTTCGACCAATGTC 6780
Db 7061 GTGACTCGATAGAGTCTGCGGTCCAAAGTGTGCTTCTGAGACTCTTCGACCAATGTC 7120
OY 6781 GAAGAAAGGAGCACTTGAAGCTTGAGCTTGATACATCAGATTAATGCTCCCAAGAGAG 6840
Db 7121 GAAGAAAGGAGCACTTGAAGCTTGAGCTTGATACATCAGATTAATGCTCCCAAGAGAG 7180
OY 6841 TTCCACACAGCTTTACCGGCTTGAGGACGCGCTGATTAACACCAACGCTTGTGAATG 6900
Db 7181 TTCCACACAGCTTTACCGGCTTGAGGACGCGCTGATTAACACCAACGCTTGTGAATG 7240
OY 6901 TGAAGAAAGGCAAGATTAACACCGGCACTGTTGGGGGCTGTGCTCTCTCTCTAG 6960
Db 7241 TGAAGAAAGGCAAGATTAACACCGGCACTGTTGGGGGCTGTGCTCTCTCTCTAG 7300
OY 6961 AAAACCCGACGCTCCCGCAAGAGAGCGCGGACAGTGGGCTTAAGTAGAGACTCCATA 7020
Db 7301 AAAACCCGACGCTCCCGCAAGAGAGCGCGGACAGTGGGCTTAAGTAGAGACTCCATA 7360
OY 7021 GAGATGCTCCCTTCAACAGCTGGCATTAAGTCTTTGACAGCCCCCCCAAGCGGCGAT 7080
Db 7361 GAGATGCTCCCTTCAACAGCTGGCATTAAGTCTTTGACAGCCCCCCCAAGCGGCGAT 7420
OY 7081 TCAAGCTTTTCCACGGGGGCGGGGCTGCGCATTCGCGGACATCAGAGGCTCTGTATGAG 7140
Db 7421 TCAAGCTTTTCCACGGGGGCGGGGCTGCGCATTCGCGGACATCAGAGGCTCTGTATGAG 7480
OY 7141 TTGGCCCTTTGCGAGACAGGTTTCATCTCTTCATGCCCCCTCCAGAGGGGAGCTTGA 7200
Db 7481 TTGGCCCTTTGCGAGACAGGTTTCATCTCTTCATGCCCCCTCCAGAGGGGAGCTTGA 7540
OY 7201 GATTCAGACCTGAGGCTTGAGAGAGTGAAGCCCAACCCCCCCCAAGGGGGGTGGCA 7260
Db 7541 GATTCAGACCTGAGGCTTGAGAGAGTGAAGCCCAACCCCCCCCAAGGGGGGTGGCA 7600
OY 7261 GCTCCGGGCTCGGACCTCGGGGCTCTGCTCATCTTGCTCGAGAGAGACCATCCGTCGTG 7320
Db 7601 GCTCCGGGCTCGGACCTCGGGGCTCTGCTCATCTTGCTCGAGAGAGACCATCCGTCGTG 7660
OY 7321 TGCTGCTCATGTCATATCTCTGGAACGGGGCTCTAATACTCTTGTAGTCCGAGAG 7380
Db 7661 TGCTGCTCATGTCATATCTCTGGAACGGGGCTCTAATACTCTTGTAGTCCGAGAG 7720
OY 7381 GAGAAGTTACCGATTAACCCCTTGAGCACTCCCTGTGGATATCAACAAGGTGTAC 7440
Db 7721 GAGAAGTTACCGATTAACCCCTTGAGCACTCCCTGTGGATATCAACAAGGTGTAC 7780
OY 7441 TGTACCAACAAGAGCGGCTCACTAAGGCTTAAAGGTAACTTTGATAGGATGCA 7500
Db 7781 TGTACCAACAAGAGCGGCTCACTAAGGCTTAAAGGTAACTTTGATAGGATGCA 7840
OY 7501 GTGCTGACTCTACTACGACTCAGTCTTAAAGGACATTAAGCTAGCGGCTCAAGGTC 7560
Db 7841 GTGCTGACTCTACTACGACTCAGTCTTAAAGGACATTAAGCTAGCGGCTCAAGGTC 7900
OY 7561 ACCGCAAGGCTCTTCAACATGAGAGAGGCTTGCAAGTAAACCCCACTTCTGCAAG 7620
Db 7901 ACCGCAAGGCTCTTCAACATGAGAGAGGCTTGCAAGTAAACCCCACTTCTGCAAG 7960
OY 7621 TCTTAAATATGAGGTTTGGGCTTAAGAGAGTCCGAGGCTTGCGGGAGGGCCGTTAACAC 7680
Db 7961 TCTTAAATATGAGGTTTGGGCTTAAGAGAGTCCGAGGCTTGCGGGAGGGCCGTTAACAC 8020
OY 7681 ATCAAGTCCGCTGGAAGAGCTCTCGAGAGACTCAGAAACACCAATTTCCCAACATTT 7740
Db 8021 ATCAAGTCCGCTGGAAGAGCTCTCGAGAGACTCAGAAACACCAATTTCCCAACATTT 8080

QY 7741 ATGGCCAAAATAGAGTGTCTGCGTGAACCCCAACAAAGGGGCAAGAAAGAGCTGCG 7800
 Db 8081 ATGGCCAAAATAGAGTGTCTGCGTGAACCCCAACAAAGGGGCAAGAAAGAGCTGCG 8140
 QY 7801 CTTATCGTTTAACTGACCTGCGGCTGAGGGGTCTGCGAAGATGCGCTTTATGACATT 7860
 Db 8141 CTTATCGTTTAACTGACCTGCGGCTGAGGGGTCTGCGAAGATGCGCTTTATGACATT 8200
 QY 7861 ACACAAAATTTCTGACGGGCGTGAAGGGGCTTTATGATTCAGTATCCCGCT 7920
 Db 8201 ACACAAAATTTCTGACGGGCGTGAAGGGGCTTTATGATTCAGTATCCCGCT 8260
 QY 7921 CACGGGTAGAGTTTCTTGAAGCATGAGCGGAAAGAGACCTTATGAGTTTTCG 7980
 Db 8261 CACGGGTAGAGTTTCTTGAAGCATGAGCGGAAAGAGACCTTATGAGTTTTCG 8320
 QY 7981 TATGATACCCGATGCTTGAATCAACCGTCACTGAGAGACATCAGAGTGAAGTCC 8040
 Db 8321 TATGATACCCGATGCTTGAATCAACCGTCACTGAGAGACATCAGAGTGAAGTCC 8380
 QY 8041 ATATATCGGGCTGCTCTTGGCCGAGAGAGCCCACTGCCATACCTGCTTAACTGAG 8100
 Db 8381 ATATATCGGGCTGCTCTTGGCCGAGAGAGCCCACTGCCATACCTGCTTAACTGAG 8440
 QY 8101 AGACTTTACGTGAGAGGGCTTATGTTCAACAGCAAGGGCCAAACCTGCGGTTACAGCGT 8160
 Db 8441 AGACTTTACGTGAGAGGGCTTATGTTCAACAGCAAGGGCCAAACCTGCGGTTACAGCGT 8500
 QY 8161 TGGCGGCGCAGCGGGGTGCTCACCACTAGCATGAGGAAACCATCAGACTTACGTGAA 8220
 Db 8501 TGGCGGCGCAGCGGGGTGCTCACCACTAGCATGAGGAAACCATCAGACTTACGTGAA 8560
 QY 8221 GCGTTAGCGGCTTGAAGTGAAGGATTAATGCGGCCCAATGCTGATATGCGGGAT 8280
 Db 8561 GCGTTAGCGGCTTGAAGTGAAGGATTAATGCGGCCCAATGCTGATATGCGGGAT 8620
 QY 8281 GACTTGGTTGATCTCAGAAAGCAGAGGAGCAGAGAGAGCAGAGGAACTTGAAGGCC 8340
 Db 8621 GACTTGGTTGATCTCAGAAAGCAGAGGAGCAGAGAGAGCAGAGGAACTTGAAGGCC 8680
 QY 8341 TTGACGAGAGCTATGACAGAGTATTCGCTCTGCTGAGACCCCGCAGACCGAGTAT 8400
 Db 8681 TTGACGAGAGCTATGACAGAGTATTCGCTCTGCTGAGACCCCGCAGACCGAGTAT 8740
 QY 8401 GATCTGAGAGCTGATACATCTTGTCTCTCAATGCTGTGCGGCTGCGGCCCAAGAGC 8460
 Db 8741 GATCTGAGAGCTGATACATCTTGTCTCTCAATGCTGTGCGGCTGCGGCCCAAGAGC 8800
 QY 8461 CGCGGAGATACCTGACAGAGACCTTACCACTCAATGCGCGGCGTGCCTGGAA 8520
 Db 8801 CGCGGAGATACCTGACAGAGACCTTACCACTCAATGCGCGGCGTGCCTGGAA 8860
 QY 8521 ACAGTTAGACATCCCTGTCTCAATTCATGCTGAGGAAACATTCAGTACGCGCCGAC 8580
 Db 8861 ACAGTTAGACATCCCTGTCTCAATTCATGCTGAGGAAACATTCAGTACGCGCCGAC 8920
 QY 8581 ATATGCGCTGCAATGCTGCTGATGACACATCTTCTTCTCAATTCATGCTGCAAGACG 8640
 Db 8921 ATATGCGCTGCAATGCTGCTGATGACACATCTTCTTCTCAATTCATGCTGCAAGACG 8980
 QY 8641 CTGAGACAGAACTCAACTTGAAGTATGAGAGCGGTGACTCCGAGGCTCCCTTGAG 8700
 Db 8981 CTGAGACAGAACTCAACTTGAAGTATGAGAGCGGTGACTCCGAGGCTCCCTTGAG 9040
 QY 8701 CTGAGACAGAACTCAACTTGAAGTATGAGAGCGGTGACTCCCTTGAGCAACATCACT 8760
 Db 9041 CTGAGACAGAACTCAACTTGAAGTATGAGAGCGGTGACTCCCTTGAGCAACATCACT 9100
 QY 8761 CCCCAGAACTGACAGGGGCTTCAAGCCCTCAAGAAACTTGGGGGCGCAACCCCTGAG 8820
 Db 9101 CCCCAGAACTGACAGGGGCTTCAAGCCCTCAAGAAACTTGGGGGCGCAACCCCTGAG 9160
 QY 8821 GCGTGAAGAGCGGGGACGTGAGTCAAGGCGTCCCTCATCTCCGTGGGGGAGAGCG 8880

Db 9161 GCGTGAAGAGCGGGGACGTGACGTACAGGCGCTCCCTCATCTCCGTGGGGGAGAGCG 9220
 QY 8881 GCCGTTGCGGCTGATATCTTTCATTAATGGGCGGTGAAGCAAGCTCAAACTCACTCCA 8940
 Db 9221 GCCGTTGCGGCTGATATCTTTCATTAATGGGCGGTGAAGCAAGCTCAAACTCACTCCA 9280
 QY 8941 TTGCGGGAAGCGGCGCTCCGTGATTTATCCAGCTGGTTCACCGTCCGCGCGGGGAGC 9000
 Db 9281 TTGCGGGAAGCGGCGCTCCGTGATTTATCCAGCTGGTTCACCGTCCGCGCGGGGAGC 9340
 QY 9001 GACATTTATCAGACGCTGTCGCTGCGCCGACCCCGCTTATGCTTTTGCTTACTCTTA 9060
 Db 9341 GACATTTATCAGACGCTGTCGCTGCGCCGACCCCGCTTATGCTTTTGCTTACTCTTA 9400
 QY 9061 CTTTGTGAGGGTGAAGCTTTTCTTACTTCCCGCTCGG 9099
 Db 9401 CTTTGTGAGGGTGAAGCTTTTCTTACTTCCCGCTCGG 9439

RESULT 5

HEPOLP 9589 bp RNA linear VRL 20-JUN-1998
 LOCUS Hepatitis C virus genomic RNA for polyprotein, complete cds.
 DEFINITION D00944
 ACCESSION D00944.1 GI:221650
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Hepatitis C virus
 Hepatitis C virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE

1 (bases 1 to 9589)
 AUTHORS Okamoto, H., Okada, S., Sugiyama, Y., Kurai, K., Iizuka, H., Machida, A., Miyakawa, Y., and Kayama, M.
 TITLE Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions
 JOURNAL J. Gen. Virol. 72 (Pt 11), 2697-2704 (1991)
 MEDLINE 92044440
 PUBMED 1658196

REFERENCE

2 (bases)
 AUTHORS Han, J. H. and Houghton, M.
 TITLE Group specific sequences and conserved secondary structures at the 3' end of HCV genome and its implication for viral replication
 JOURNAL Nucleic Acids Res. 20 (13), 3520 (1992)
 MEDLINE 92335016
 PUBMED 1321416

REFERENCE

3 (bases)
 AUTHORS Hotta, H., Doi, H., Hayashi, T., Purwanta, M., Soemarto, W., Mizokami, M., Ohba, K. and Homma, M.
 TITLE Analysis of the core and E1 envelope region sequences of a novel variant of hepatitis C virus obtained in Indonesia
 JOURNAL Arch. Virol. 136 (1-2), 53-62 (1994)
 MEDLINE 94270990
 PUBMED 7545932

COMMENT

These data kindly submitted in computer readable form by: Hiroaki

FEATURES

source

1. 9589
 Location/Qualifiers
 /organism="Hepatitis C virus"
 /mol_type="genomic RNA"
 /isolate="HC-J6"
 /db_xref="taxon:11103"
 341..9442
 /codon_start=1
 /product="polyprotein"

/protein_id="BAA00792.1"
/db_xref="gi:221651"
/translation="WSTNPKPQRTKRTNTRRPODVKPRGGGQIVGGVLLPRGRPL
GVRATRKTSERSSOPGRROPIDKDRSGTGKMGKPGVPMPIYGNBGIAMALHAGRVLED
SRPSPGPNDRPBRNRNOKVIDITLCGRADIMGTIPVVGALGVNABALHAGRVLED
GVNFRATGMPCGSFSLFLLALLSCITTVSAEYKNISTGTMTNDCTNDISTIMQOLA
AVLHVGPCVCEKGNSTRSMIPVSPNVAQOPGALTOGLTHIDMVMNMTCSALY
VGLDCCGVMLAOMFIVSPQHMFEVDNCISIYPGTTIGHMADMMNNSPTATMPL
AYAMRPEVVIDIIGAHMGVFMGLAFPMOGAMAKVVILLLAGDAQHTVGGST
AHNATLTGMSFSLGRKQIOLINTNGSWHNRALNCDSLHTGFLSLFTFHSFNS
GCPRMSCRSITEAPRWGMGALOYEDNTPEDMRPCMHPRPCCGVASVSSGVCV
YCFPRPVYVGTTRDLGAPITYTNGENEDVFLNSTRPQSMNGCTMAMSTGYTKC
GAPCRIRADFNASMDLCTPDCERKHPDITYIKCGSPMLTPCLIDYPRLMHYEC
TVNVTIFKIRMYGVGVEHRITPAACNFRGDRNCLEDRSQSLSTGLTSLTMMALLPT
YSDLPALSTGLLHONIVDOFMGSLPALTKYIVMEVLLFLLLAARVACACM
MLILLOAEALAEKLVLAASACNGLFEVIFPVAAYIKRVAWLYTSITGLM
SGLILLALPOOAVYDASVHGOIGALLVITLPTLTPOYKTLSPFLMWICVLTLL
ABAMVQWAPPMQVRGRGDITMAVAIFCPGVVDITKMLAVGPRVILKGLTRIP
YFVRHALHMCNTRHLAGRVYOMVVLALGRNTGYITDHLTPMSDMANGLDLA
VAVEPIISPMKEKVIWGAETAACGDLHGLPVASRGRVLLGPDADYVTSKMSLL
APITVAQOTGILGTTIVSMTGDKTEQAGEIOVLSTVTSPLGTTISGLMTVYHG
AGNKTLAGSRGVITQWSSASGDLGVMSPGTSLBECTCGAIDLILNRNAVITA
RRGDKGALLSPRLSTLKGSSGGPVLCPRGAVGVRAVCRCVAKSIDIPVET
LDITVRSPFSDNSTPPAVPOTYGVYHAPITGSGKSTKVAVAAGYKVLNPSV
AATLDFGAYLSKAGHINPIRTGARTVTGAPITTYTGKFLADGGAGAYDIIICD
ECHANVSTIIIGITVLDQASTAGVRLVLTATPPGCVTTPHNIEVALGQGEIP
FYGRAIPLSYIKGGRHLIFCHSKKCDLALALRMGMANAVAYRGDLVSIPTQGV
VVVATDALMTGTGDFSDVIDCNVAIVQVDFSLDPTETITPTQVPODASRGRBR
TGRGLGTYRVYVSTGERBASGMFDSVLECCDACAAMVELPATTTIRLAPYRTPPL
PVCODHLEFMEAVFTGLTHIDHFLSGTQSGENPAYITAOATVCAKARPPSMV
MMKCLTRLPKPLVGPPLLYRLGVSNTVNLTHPTVKIATACODALEMTVWGLG
GMLAAVAVYCLATGCVCIIIGRLHVNQRAVVAIPDKVLEADENBEMASRALIIEBO
RIAEMLKSKIOGLLOQASQKQADIDPAVQASMPYKQWAGMNNPISGLOGLAST
LPBNPAVSMMAFSALTSPLSTSTTLINLTGMLASOILPAGAIFFVYGLVGA
VCSIGLKVLYVDIILAGYAGISGALPAKIMSGERKSEDEVDNLPGLISGLVYV
ICAILLRHVGBGAVVMMRLIAPSRGNVAPTHYVTSVDSLSQVTOGLSLTIT
SLRLHNMVITBDCPIPCSGSMRLDVMWVCTIILDFENMLTSLPLPMPLPTISQ
KGYKVMAGTGMITRCPGANISGNVAIGSMRTGPTCNMIQOGFIPNICYBDC
VPRAPNFKAIMRWASBYAEVTOHSGYHITGLTDTNLVPCQGLSPFSSVYDQ
OIHRAPIPKPEPDEVFCVGLNFFVGSOLPCDPREDTDLVLSMLDPIHIAETA
ABRLARGSPSEBASASOLSAPSLRACCTTHGKAYIDVMDALPMGQDPILES
KYVVDLSIDPWEERSDLBESTPSEYMLPKKFPALPAMKRPYNNPLVDSMRPY
OPATVAGCALPPKPTTPPPRRRTVPLSSSISGADALQOALIFESQPPGSGSIS
TGADADSGSRTPPDELASETGSISMPLEGGEDDLDBOVLBPPOGCVNTP
GSGSGMSSTCEEDSVVCSMSYMTGALITPCSPBEKPIINPLNSILRYNKY
CTTSKASLRAKVTFDROMALDAHYDVKDILKLAASKVATRLTIEBAOLPPHS
ASRYGPAKVRSLSGRAVNIKSVMDLLEDPOTPIPTIMAKNVPCDPKRGK
KAARLIVPDLQVRCERKALYDITOKIPOAVMAGSTFOYSPARVEFLKAMAEK
DPMGSDYTRCFDSTVTERDIRTESIYRACSLPEBARTALHSILERYVGGPMPNK
GQTCGRCRASGVLITSMGNTITCYVVALACRAAGIAPTMVCGDILVISESG
TEEDBRNLRAFTBAMTRYAPGDPREYDLALITSCSVMSVALGQGRRYLYLR
DPTPIIARAAMETVHSPVNSWMLGNIIOYAPTIWARMHPSILMAOQTLPAWS
PEMGAIVSVSRPLDPAIIERLHGLDASLHTYRPHETLRVASLRLKIGAPLPAWS
RARAVRASLISRGRAVCGRYLPMAVKTKLTPLPEARLLDLSHFTVAGSGDI
YHSVRARPRLLGLLFLVGVGLFLPAR"

ORIGIN

Query Match 93.4%; Score 8497.4; DB 14; Length 9589;

Beef Local Similarity 95.9%; Pred. No. 0; Mismatches 376; Indels 0; Gaps 0;

Matches 8723; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

QY 1 ATGAGACAAATCTTAAACCTCAAGAAAAACCAAAAGAAACCAACCAACCGTGCACCA 60
DB 341 ATGAGACAAATCTTAAACCTCAAGAAAAACCAAAAGAAACCAACCAACCGTGCACCA 400
QY 61 GACGTTAAGTTCCGGGCGGCGGCGGCAATCTTGGCGGAGTACTTGTGGCGGCGAGG 120
DB 401 GACGTTAAGTTCCGGGCGGCGGCGGCAATCTTGGCGGAGTACTTGTGGCGGCGAGG 460
QY 121 GGGCCCAAGGTTGGGTGTGGCGGCAAGAAAGAACTTGGAGCGGTCCAGCGCACTGGA 180
DB 461 GGGCCCAAGGTTGGGTGTGGCGGCAAGAAAGAACTTGGAGCGGTCCAGCGCACTGGA 520

QY 181 AGGCGCCAGCCCATCCCTAAAGATCGGCGCTCCATCGGCAATCTTGGGAAAAACCAAGA 240
DB 521 AGGCGCCAGCCCATCCCTAAAGATCGGCGCTCCATCGGCAATCTTGGGAAAAACCAAGA 580
QY 241 TACCCCTGGCCCTTATACGGAAATGAGGACTTCGCTGGGACAGATGGCTCTGTCCCCC 300
DB 581 TACCCCTGGCCCTTATACGGAAATGAGGACTTCGCTGGGACAGATGGCTCTGTCCCCC 640
QY 301 CGAGTTCCCGTCCCTCTTGGGAGCCCAATGACCCCGGACATAGGTGCGGCAACGTGGT 360
DB 641 CGAGTTCCCGTCCCTCTTGGGAGCCCAATGACCCCGGACATAGGTGCGGCAACGTGGT 700
QY 361 AAGTCATGATACCTTAACGTGCGGCTTGCACCTTAAGGGATACCTCCTGTCTG 420
DB 701 AAGTCATGATACCTTAACGTGCGGCTTGCACCTTAAGGGATACCTCCTGTCTG 760
QY 421 GGGGCCCGCTGGGCGGCGCTCGCAAGCTTCGCGGCAAGGGTGAATCCTCGAGAGAC 480
DB 761 GGGGCCCGCTGGGCGGCGCTCGCAAGCTTCGCGGCAAGGGTGAATCCTCGAGAGAC 820
QY 481 GGGGTTAATTTTGAACAAGGAACTTACCCGTTGCTCTTTCTATCTTCTTGCTGACC 540
DB 821 GGGGTTAATTTTGAACAAGGAACTTACCCGTTGCTCTTTCTATCTTCTTGCTGACC 880
QY 541 CTGCTGTCTGATCAACAACCCCGGTCTCTCGCTGCGGAAGTGAAGAACTACGTACGCG 600
DB 881 CTGCTGTCTGATCAACAACCCCGGTCTCTCGCTGCGGAAGTGAAGAACTACGTACGCG 940
QY 601 TACATGATGATCAAGACGTGCACCAATGACAGATTAAGTGCAGTCCAGGCTGTGTC 660
DB 941 TACATGATGATCAAGACGTGCACCAATGACAGATTAAGTGCAGTCCAGGCTGTGTC 1000
QY 661 CTCACAGTCCCGGAGTGCCTCCGTCGAGAAAGTGGGAAATGCAATCTCAGTGTGATA 720
DB 1001 CTCACAGTCCCGGAGTGCCTCCGTCGAGAAAGTGGGAAATGCAATCTCAGTGTGATA 1060
QY 721 CCGGTCTACCGAATGTGCGCGGTGACAGGAGCCGCGCTTACGAGGCTTGGAGAG 780
DB 1061 CCGGTCTACCGAATGTGCGCGGTGACAGGAGCCGCGCTTACGAGGCTTGGAGAG 1120
QY 781 CACATGACATGATTTGATGTCGCGCACGCTGTGCTCGCTTACGAGGAGGAGCTTC 840
DB 1121 CACATGACATGATTTGATGTCGCGCACGCTGTGCTCGCTTACGAGGAGGAGCTTC 1180
QY 841 TGGGTGGGATGATGCTCCAGCCCAATGTTCAATTTGTCCGCGGAGACCACTGTTT 900
DB 1181 TGGGTGGGATGATGCTCCAGCCCAATGTTCAATTTGTCCGCGGAGACCACTGTTT 1240
QY 901 GTCCAAAGCTGCAATTTGCTTCATCTACCTGTGATCACTGACACCGCATGCGATGG 960
DB 1241 GTCCAAAGCTGCAATTTGCTTCATCTACCTGTGATCACTGACACCGCATGCGATGG 1300
QY 961 GACATGATGATGAACTGTGTCGCGCAAGGCTACATGATGCGGAGAGTGGCGTGC 1020
DB 1301 GACATGATGATGAACTGTGTCGCGCAAGGCTACATGATGCGGAGAGTGGCGTGC 1360
QY 1021 CCGGAGTCAATTATAGACATCATTTAGCGGAGCTCATTTGGGAGCTCATGTTGCGCTTGGCC 1080
DB 1361 CCGGAGTCAATTATAGACATCATTTAGCGGAGCTCATTTGGGAGCTCATGTTGCGCTTGGCC 1420
QY 1081 TACTTCTTATGCAAGGAGCGTGGCGAAAGTCTGTTGTCATCTTGTGGCGCGGAG 1140
DB 1421 TACTTCTTATGCAAGGAGCGTGGCGAAAGTCTGTTGTCATCTTGTGGCGCGGAG 1480
QY 1141 GTGAGCGCGGCAACCATCTGTGTGGGGGTTCTGCGGCGAGACCAACCGGAGCGCTACCC 1200
DB 1481 GTGAGCGCGGCAACCATCTGTGTGGGGGTTCTACCGCGATTAACCCAGAACCTTACCC 1540
QY 1201 AGCTATTTGACATGGGCGCCCAAGGCAAAAATCAGCTGTTTAAACCAATGGCAGCTGG 1260
DB 1541 GGCATGTTCTCCTTGTGTGCGAGGCAAGAAATCAGCTATCAACCAATGGCAGCTGG 1600
QY 1261 CACATCAACCGCACCGCGCTGAACCTGCAATGACTCTTGTGACACCGGCTTTATGCGCTCT 1320

Db 1601 CAGATCAACCGGCAACCGCCCTGMACTGCAAGACTCTTTGCAACACCGGCTTCCTGCTCA 1660
Qy 1321 CTGTTCTACACCAACAGCTTCAACTGTCAGAGATGTCGCCAGCGATGTCGCGCTGCGC 1380
Db 1661 CTCTTCTACACCAACAGCTTCAACTGTCAGAGATGTCGCCAGCGATGTCGCGCTGCGC 1720
Qy 1381 AGTATGAGAGCCTTCCGGGTGGAGATGGGGGGCCCTGGAATATAGAGATATATGCAACAT 1440
Db 1721 AGATGAGAGCCTTCCGGGTGGAGATGGGGGGCCCTGGAATATAGAGATATATGCAACAT 1780
Qy 1441 CCAGAGATATAGAACCCCTATTTGCTGCACTACCCACCAAGAGAGCTGTCGCTGCC 1500
Db 1781 CCAGAGATATAGAACCCCTATTTGCTGCACTACCCACCAAGAGAGCTGTCGCTGCC 1840
Qy 1501 GCCAAGACTGTGTGGCCCAAGTGTACTGTTTCAACCCCAAGCCCAAGTGTGTGGCAGC 1560
Db 1841 GCGAGCTGTGTGGCCCAAGTGTACTGTTTCAACCCCAAGCCCAAGTGTGTGGCAGC 1900
Qy 1561 ACCGACAGGCTTGGAGGGCCCACTTACAGTGGGGGGAGATGAGACAGATGCTTCTCA 1620
Db 1901 ACCGATAGACTTGGAGGGCCCACTTACAGTGGGGGGAGATGAGACAGATGCTTCTCA 1960
Qy 1621 TTGAAGAGCACTGACCAACCGCTGGGGTCAATGGTTCCGCTGCACTGAGATGAATCTCT 1680
Db 1961 TTGAAGAGCACTGACCAACCGGAGGGTCAATGGTTCCGCTGCACTGAGATGAATCTCA 2020
Qy 1681 GCGTACCAAGACTTGGGGGCAACACCTGCGCTATCTAGAGCTGACTTCAACGCCAGC 1740
Db 2021 GCGTACCAAGACTTGGGGGCAACACCTGCGCTATCTAGAGCTGACTTCAATGCGCAGC 2080
Qy 1741 ACGGACCTGTGTGGCCCAAGTGTATGAGAGATGCTATACCACTTACCTCAAA 1800
Db 2081 ATGGAATGTTGTGGCCCAAGTGTATGAGAGATGCTATACCACTTACCTCAAA 2140
Qy 1801 TCGGCTCTGGGGCCCTGCTGCTCAACGCAAGGTGCTGATCACTAACCCCTACAGGCTTGG 1860
Db 2141 TGTGGCTCTGGGGCCCTGCTGCTCAACGCAAGGTGCTGATCACTAACCCCTACAGGCTTGG 2200
Qy 1861 CATTACCCCTGCAACAGTTAACTATACCATCTTCAAAATAGATATGTGGAGGGGTT 1920
Db 2201 CATTACCCCTGCAACAGTTAACTATACCATCTTCAAAATAGATATGTGGAGGGGTT 2260
Qy 1921 GACACAGAGCTCAACGCTGATGCAATTTAGATCTGAGGGAGATGTTGCACTTGGAGAGC 1980
Db 2261 GACACAGAGCTCAACGCTGATGCAATTTAGATCTGAGGGAGATGTTGCACTTGGAGAGC 2320
Qy 1981 AGAGACAGAGTCAACTGTCTCTCTTGTGTCACCTCCACAGAGATGGGCAATTTTACT 2040
Db 2321 AGAGACAGAGTCAACTGTCTCTCTTGTGTCACCTCCACAGAGATGGGCAATTTTACT 2380
Qy 2041 TGTCTTTTACTCGGACCTGCGCGCTTGTGCACTGTGTCTTCCACTCCACCAAAACATC 2100
Db 2381 TGTCTTTTACTCGGACCTGCGCGCTTGTGCACTGTGTCTTCCACTCCACCAAAACATC 2440
Qy 2101 GTGGAAGTCAATTCATGTATGAGCTATCACTGCGCTCAAAATATACATGTCGCAATG 2160
Db 2441 GTGGAAGTCAATTCATGTATGAGCTATCACTGCGCTCAAAATATACATGTCGCAATG 2500
Qy 2161 GAGTGGTAACTACTTATTCCTGCTCTTACGCGAGCCAGGGTGTGCGCTCTTATG 2220
Db 2501 GAGTGGTAACTACTTATTCCTGCTCTTACGCGAGCCAGGGTGTGCGCTCTTATG 2560
Qy 2221 ATGCTCATCTTGTGGGCGCAGGCGCAAGACAGCACTAGAGAGCTGTCTTGGCAGCT 2280
Db 2561 ATGCTCATCTTGTGGGCGCAGGCGCAAGACAGCACTAGAGAGCTGTCTTGGCAGCT 2620
Qy 2281 GCGAGGCGAGCTAGCTGCAATGGCTCTATATTTTGTATCTTTTGTGGGCGCTTGG 2340
Db 2621 GCGAGGCGAGCTAGCTGCAATGGCTCTATATTTTGTATCTTTTGTGGGCGCTTGG 2680
Qy 2341 TACATCAAGGCTGGGTAGTCCCTTAGCTATCTATCTCTCACTGGCTGTGGTCTTT 2400

Db 2681 TACATCAAGGCTGGGTAGTCCCTTAGCTATCTATCTCTCACTGGCTGTGGTCTTT 2740
Qy 2401 AGCTTACTGCTCTTACGATTTGCCCCAACAGGCTTATGCTTATGACGATGTGATGTC 2460
Db 2741 GGCCTACTGCTCTTACGATTTGCCCCAACAGGCTTATGCTTATGACGATGTGATGTC 2800
Qy 2461 CAGATGAGAGCGGCTGTGTGATATGATGATCACTCTCTTATCTCTCACCCCCGGTATAG 2520
Db 2801 CAGATGAGAGCGGCTGTGTGATATGATGATCACTCTCTTATCTCTCACCCCCGGTATAG 2860
Qy 2521 ACCCTTCTACGCGGTTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
Db 2861 ACCCTTCTACGCGGTTTTTGT 2920
Qy 2581 GTCCAGAGAGTGGGACCACTATATGACAGTGTGCGGGGGCCGTATGTCATATATGGGCC 2640
Db 2921 GTCCAGAGAGTGGGACCACTATATGACAGTGTGCGGGGGCCGTATGTCATATATGGGCC 2980
Qy 2641 GTCCCAATATTTACACGAGT 2700
Db 2981 GTCCCAATATTTACACGAGT 3040
Qy 2701 GGGCTGTCTTACCTCTTAAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2760
Db 3041 GGGCTGTCTTATCTCTTAAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3100
Qy 2761 GCTCTACTAGAGATGT 2820
Db 3101 GCTCTACTAGAGATGT 3160
Qy 2821 GCGCTACTAGCCCTTGGAGAGT 2880
Db 3161 GTCCTACTAGCCCTTGGAGAGT 3220
Qy 2881 TCGGATTTGGGCTGT 2940
Db 3221 TCGGATTTGGGCTGT 3280
Qy 2941 AGTCCGATGGAGAAAGTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3000
Db 3281 AGTCCGATGGAGAAAGTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3340
Qy 3001 TTACAGGACCTTCCGCTGT 3060
Db 3341 TTACAGGACCTTCCGCTGT 3400
Qy 3061 GGCCTATACCTCCAAAGGGT 3120
Db 3401 GGCCTATACCTCCAAAGGGT 3460
Qy 3121 CGTGGCTTTTGGGCACTATGT 3180
Db 3461 CCGGGCTTTTGGGCACTATGT 3520
Qy 3181 GGGGAATTCAGGCTCTGT 3240
Db 3521 GGGGAATTCAGGCTCTGT 3580
Qy 3241 GTTTTGTGACGT 3300
Db 3581 GTTTTGTGACGT 3640
Qy 3301 GTTACGCAATGTATCTTCAAGT 3360
Db 3641 GTTACGCAATGTATCTTCAAGT 3700
Qy 3361 ACTTAAATCTTTGAGGCGGT 3420
Db 3701 ACTTAAATCTTTGAGGCGGT 3760
Qy 3421 GCTGATGTATCTCCGCTGT 3480
Db 3761 GCTGATGTATCTCCGCTGT 3820

Qy 3481 CCTTTTCAACCTTGAAGGGGTCTCTCAGAGAGCCCGGTGCTATGCCCCAGAGGGCCACGCT 3540
Db 3821 CCTTTTTCACCTTGAAGGGGTCTCTCGAGGGGGCCCGGTGCTGCTGCCCAAGAGCCACGCT 3880
Qy 3541 GTCGAGCTCTTCCGAGGACGCTGTGTGCTCTCGAGGGCTGTGCTAAGTCCATATGATTTTCATC 3600
Db 3881 GTCGAGGCTCTTCCGAGGACGCTGTGTGCTCTCGAGGGCTGTGCTAAGTCCATATGATTTTCATC 3940
Qy 3601 CCGGTGAGACACTGACATGCTGAGGAGGTCCCGCACTTTAGTAGCAACAGACACCA 3660
Db 3941 CCGGTGAGACACTGAGATGCTGACTGCTGCTCCCGCACTTTAGTAGCAACAGACACCA 4000
Qy 3661 CTTGCTGTGCCCCAGACCTTATCAGGTGCGGTACTTGTGATGCCCCAGCTGGAGTGAAG 3720
Db 4001 CTTGCTGTGCCCCAAGCTTATCAGGTGCGGTACTTATGATGCCCCAGCTGGAGTGAAG 4060
Qy 3721 AGCACAAGTCTCTGTGCTATGCTGTCTCAGAGGGTATTAAGTGTCTAATGCC 3780
Db 4061 AGCACAAGTCTCTGTGCTATGCTGTCTCAGAGGGTATTAAGTGTCTAATGCC 4120
Qy 3781 TCAGTGTGCTGACCTGTGAGGTTTGGAGGCTTGTCTAAGGACATGAGCATCAATCCC 3840
Db 4121 TCGGTGCTGCTGACCTGTGAGGTTTGGAGGCTTGTCTAAGGACATGAGCATCAATCCC 4180
Qy 3841 AACATTAGAGCTGAGAGTCAAGAGCTGTGACGACCGAGGAGCCCATCAGCTACTCCACATAT 3900
Db 4181 AACATTAGAGCTGAGGAGTCAAGAGCTGTGACGACCGAGGAGCCCATCAGCTACTCCACATAT 4240
Qy 3901 GGCAAATTCCTGCGCCGATGAGGGGGCTGTGCGAGGGCCGCTTACGACATCATATGATGAT 3960
Db 4241 GGCAAATTCCTGCGCCGATGAGGGGGCTGTGCGAGGGCCGCTTATGACATCATATGATGAT 4300
Qy 3961 GAATGCACTGCGAGTCTTACCAACCATCTTGTGACATCGGAAAGTCTTATCAAGCA 4020
Db 4301 GAATGCACTGCGAGTCTTACCAACCATCTTGTGACATCGGAAAGTCTTATCAAGCA 4360
Qy 4021 GAGACAGCTGAGGCTGAGACTGTGCTGCTACAGTCAAGCCCTTGGGTCAAGTACA 4080
Db 4361 GAGACAGCTGAGGCTGAGACTGTGCTGCTACAGTCAAGCCCTTGGGTCAAGTACA 4420
Qy 4081 ACCCCCCACCCCAACATAGAGAGGTGGCCCTTGGGACAGAGGGCGAGATCCCTTCTAT 4140
Db 4421 ACCCCCCACCCCAACATAGAGAGGTGGCCCTTGGGACAGAGGGGTGAGATCCCTTCTAT 4480
Qy 4141 GGGAGGGGAGTTCCTGCTCTTACATCAAGGAGGAAAGACATCTGATCTTCTGCACTTCA 4200
Db 4481 GGGAGGGGAGTTCCTGCTCTTACATCAAGGAGGAAAGACATCTGATCTTCTGCACTTCA 4540
Qy 4201 AAGAAAAAGTGAAGAGCTGCGAGCGGCCCTTGGGGTATGGCTTGAACCTCAGTGGCA 4260
Db 4541 AAGAAAAAGTGAAGAGCTGCGAGCGGCCCTTGGGGTATGGCTTGAACCTCAGTGGCA 4600
Qy 4261 TACTACAGAGGGTGTGAGAGTCTCCGCTAATACCAACTCAGGAGAGAGTATGCTGTGCC 4320
Db 4601 TACTACAGAGGGTGTGAGAGTCTCCGCTAATACCAACTCAGGAGAGAGTATGCTGTGCC 4660
Qy 4321 ACCGAGGCTCATGACAGAGTATCTGGGACTTTGACTCCGATGCTGACATGCAAGCTA 4380
Db 4661 ACCGAGGCTCATGACAGAGTATCTGGGACTTTGACTCCGATGCTGACATGCAAGCTA 4720
Qy 4381 GCGGTCACTCAAGTTGTAGACTTCAAGTTTGAACCCACATTCACCATTAACCAAGATT 4440
Db 4721 GCGGTCACTCAAGTTGTAGACTTCAAGTTTGAACCCACATTCACCATTAACCAAGACT 4780
Qy 4441 GTCCCTCAAGAGCTGTCTCAAGTATGACGAGCGCGGGGTGCGCAAGGGAGAAAGTGTG 4500
Db 4781 GTCCCTCAAGAGCTGTCTCAAGTATGACGAGCGCGGGGTGCGCAAGGGAGAAAGTGTG 4840
Qy 4501 GGCATTTATAGTATGTTTCACTGTGAGCGAGGCTCAGGAATGTTTGAACATGTATGTG 4560
Db 4841 GGCATTTATAGTATGTTTCACTGTGAGCGAGGCTCAGGAATGTTTGAACATGTATGTG 4900

Qy 4561 CTCTGTAGTGTCTACGACGACGAGGGCCGAGTATGAGCTCAACATCGAGACACACC 4620
Db 4901 CTCTGTAGTGTCTACGATGACGAGGGCCGATGTATGAGCTCAACAGCGAGACACACC 4960
Qy 4621 GTGAGGCTCAGAGGCTATTTCAACAGCGCCGGTTTGTCTGTGTGCCAAGACCATTTGAG 4680
Db 4961 GTGAGGCTCAGAGGCTATTTCAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 5020
Qy 4681 TTTTGGAGAGGAGTTTTCACCGGCTCAGACATGATGAGCCACTTCTTCCCAACA 4740
Db 5021 TTTTGGAGAGGAGTTTTCACCGGCTCAGACATGATGAGCCACTTCTTCCCAACA 5080
Qy 4741 AAGCAATCGAGGAGAAATTTCCGATCTTAAAGCTTACAGGCTACAGTGTGCTTAG 4800
Db 5081 AAGCAATCGAGGAGAAATTTCCGATCTTAAAGCTTACAGGCTTACAGTGTGCTTAG 5140
Qy 4801 GCCAAGCCCCCCCCCGTCTGTGAGAGCTGATGGAAGTGTGACTGACTCAAGGCC 4860
Db 5141 GCCAAGCCCCCCCCCGTCTGTGAGAGCTGATGGAAGTGTGACTGACTCAAGGCC 5200
Qy 4861 ACACCTGTGAGGCCACACCTCTCTGTACCGCTTGGGCTGTGTTACCAAGAGGTACAC 4920
Db 5201 ACACCTGTGAGGCCACACCTCTCTGTACCGCTTGGGCTGTGTTACCAAGAGGTACAC 5260
Qy 4921 CTCACATATCCGTGAGCAAAATACATGCGCACTGTGATCAAGACGACCTTGAAGTATG 4980
Db 5261 CTCACATATCCGTGAGCAAAATACATGCGCACTGTGATCAAGACGACCTTGAAGTATG 5320
Qy 4981 ACCAGACATGAGTCTTGTGACGAGGGAGTCTTGGAGCGCGGTGGCGGATGCTGGGG 5040
Db 5321 ACCAGACATGAGTCTTGTGACGAGGGAGTCTTGGAGCGCGGTGGCGGATGCTGGGG 5380
Qy 5041 ACCGGGTGTGTGATCATTCGCGCGCTTGCAATTAAACAGAGAGCGTCTGTGGCGCG 5100
Db 5381 ACCGGGTGTGTGATCATTCGCGCGCTTGCAATTAAACAGAGAGCGTCTGTGGCGCG 5440
Qy 5101 GACAGAAGGCTCTTATAGAGCTTTTGTATGAGATGAGAGAAATGTGCTCTTGAAGGCGCT 5160
Db 5441 GACAGAAGGCTCTTATAGAGCTTTTGTATGAGATGAGAGAAATGTGCTCTTGAAGGCGCT 5500
Qy 5161 CTGATTGAAGAGGGGAGGAGGATAGCGAGATGCTGAAGTCCAAAGTCCAAAGCTTATTTG 5220
Db 5501 CTGATTGAAGAGGGGAGGAGGATAGCGAGATGCTGAAGTCCAAAGTCCAAAGCTTATTTG 5560
Qy 5221 CAGCAAGCTTCCAAAGCTCAGACATTAACCAACCTGAGAGCTTCAATGGCCCAAG 5280
Db 5561 CAGCAAGCTTCCAAAGCTCAGACATTAACCAACCGCTGTGAGAGCTTCTTGGGCCAAG 5620
Qy 5281 GTAGAACATTTCTGGGCAACATGTGGAATCTTATAGCGGACATCCATTAACCTTCGCA 5340
Db 5621 GTAGAGCAATTTCTGGGCAACATGTGGAATCTTATAGCGGACATCCATTAACCTTCGCA 5680
Qy 5341 GGACTATCAACATGACAGGAAACCTGACAGTATGCTTCAATGATGAGCGTTCAAGTGGCGCC 5400
Db 5681 GGACTATCAACATGACAGGAAACCTGACAGTATGCTTCAATGATGAGCGTTCAAGTGGCGCC 5740
Qy 5401 CTCACAGTCCGCTGTCAACAGACACCATATCTTCAACAATTTTGGGGGGCTGAGCTA 5460
Db 5741 CTCACAGTCCGCTGTCAACAGACACCATATCTTCAACAATTTTGGGGGGCTGAGCTA 5800
Qy 5461 GCATCCCAAAATTTGACCAACCCGCGGGGGGCTACCGGCTGTGTCTCAGTGGCTGAGTGGG 5520
Db 5801 GCATCCCAAAATTTGACCAACCCGCGGGGGGCTACCGGCTGTGTCTCAGTGGCTGAGTGGG 5860
Qy 5521 GCTGCGTGTGAGGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 5580
Db 5861 GCTGCGTGTGAGGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 5920
Qy 5581 GCGGAGATTTTGGGGGCTCTGTGCTGATTCAGATCATGTGTCGAGAGAGCCCTCCATG 5640
Db 5921 GCGGAGATTTTGGGGGCTCTGTGCTGATTCAGATCATGTGTCGAGAGAGCCCTCCATG 5980
Qy 5641 GAGGATGTGCTCAACTTGTCTGCTGGAATTTCTGTCTCCGAGTCCCTTGGTATGTGGAAGTC 5700

Db 5981 GAGAGATGTGTCAACCTGCTGCTGGAAATTCGTCTCGGGGTGCTCGGTGGAGTGC 6040
Qy 5701 ATCTGGCGGGCCATTTCTGGCCCGACACGTGGGACACCGGGGAAAGCGCCCTCAATGGATG 5760
Db 6041 ATCTGGCGGGCCATTTCTGGCCCGACACGTGGGACACCGGGGAAAGCGCTGTCTCAATGGATG 6100
Qy 5761 AATGAGACTCATTTGCTTGGCTTCAGAGAAATCAAGTGGCCCACTACAGTGGAGC 5820
Db 6101 AATAGGCTCATTTGCTTGGCTTCAGAGAAATCAAGTGGCCCACTACAGTGGAGC 6160
Qy 5821 GAATCGGAGTGGTGGAGCTGGAGCCCACTACAGTGGCCCACTACAGTGGAGC 5880
Db 6161 GAATCGGAGTGGTGGAGCTGGAGCCCACTACAGTGGCCCACTACAGTGGAGC 6220
Qy 5881 CTGAGAGAGTCCACCACTGGAATTAAGAGAGCTGGCCCACTACAGTGGCCCACTACAGTGGAGC 5940
Db 6221 CTGAGAGAGTCCACCACTGGAATTAAGAGAGCTGGCCCACTACAGTGGCCCACTACAGTGGAGC 6280
Qy 5941 CTGCGGAGTGGTGGAGCTGGAGTGGAGCCCACTACAGTGGCCCACTACAGTGGAGC 6000
Db 6281 CTGCGGAGTGGTGGAGCTGGAGTGGAGCCCACTACAGTGGCCCACTACAGTGGAGC 6340
Qy 6001 TCCAAATTTATTTCCCAAGATGCGCGCTCCCTTTGCTCTCTGCAAAAAGGGGTACAG 6060
Db 6341 TCCAAATTTGTTCCCAAGATGCGCGCTCCCTTTATCTTGTCAAAAAGGGGTACAG 6400
Qy 6061 GGGGTGGGGCGCGGACCTGGACATGACCAACAGGTGCTCTTGGCGCGCAATATCTT 6120
Db 6401 GGGGTGGGGCGCGGACCTGGACATGACCAACAGGTGCTCTTGGCGCGCAATATCTT 6460
Qy 6121 GGGCAATGTCGGCTTGGGCTCATAGAAATCAAGGGGCTTAAGACCTGCAAGAAATATCTTG 6180
Db 6461 GGGCAATGTCGGCTTGGGCTCATAGAAATCAAGGGGCTTAAGACCTGCAAGAAATATCTTG 6520
Qy 6181 CAGGGGACCTTTCTCATCAATTTGTTACACGAGGGGCAAGTGGCGGAAACCGCGACCA 6240
Db 6521 CAGGGGACCTTTCTCATCAATTTGTTACACGAGGGGCAAGTGGCGGAAACCGCGACCA 6580
Qy 6241 AACTTTAAGTGGCCATCTGAGAGGTGGCGGCTCAGAGTACCGGAGGTGAGCGAGAC 6300
Db 6581 AACTTTAAGTGGCCATCTGAGAGGTGGCGGCTCAGAGTACCGGAGGTGAGCGAGAC 6640
Qy 6301 GGGTCATACCACTAACATTAAGAGCTCAACATGATTAATGAAAGTCCCTGCGCACTA 6360
Db 6641 GGGTCATACCACTAACATTAAGAGCTCAACATGATTAATGAAAGTCCCTGCGCACTA 6700
Qy 6361 CCCTCTCCGAGTCTTTCTGAGGTGAGCGAGTGCAGATCCATAGTTTGCCTCCACA 6420
Db 6701 CCCTCTCCGAGTCTTTCTGAGGTGAGCGAGTGCAGATCCATAGTTTGCCTCCACA 6760
Qy 6421 CCGAAGCGTTTTCGGGAGTGAAGTCTGCTTGGGCTTAAATTCATTTGTCGTC 6480
Db 6761 CCGAAGCGTTTTCGGGAGTGAAGTCTGCTTGGGCTTAAATTCATTTGTCGTC 6820
Qy 6481 GGGTCCAGTTCCTTGGCGACCTGAAACCGGACAGAGCTTAATGATGTCATGCTTAC 6540
Db 6821 GGGTCCAGTTCCTTGGCGACCTGAAACCGGACAGAGCTTAATGATGTCATGCTTAC 6880
Qy 6541 GATCCTCTCATATCAACGCGAGAGCTGACGCGGCTTAAAGCGGGGTACACCCCA 6600
Db 6881 GATCCTCTCATATCAACGCGAGAGCTGACGCGGCTTAAAGCGGGGTACACCCCA 6940
Qy 6601 TCCGAGGAGAGTCTGCGGAGAGCTGACGCGGCTTAAAGCGGGGTACACCCCA 6660
Db 6941 TCCGAGGAGAGTCTGCGGAGAGCTGACGCGGCTTAAAGCGGGGTACACCCCA 7000
Qy 6661 ACCCAGGCAAGAGCTTATGATGAGACATGATGGATGCTTAAACCTGTTCAATGGGGGGAT 6720
Db 7001 ACCCAGGCAAGAGCTTATGATGAGACATGATGGATGCTTAAACCTGTTCAATGGGGGGAT 7060
Qy 6721 GTGACTCGATAGAGTCTGGGCTCAAGGCTGCTTGAAGCTTCTGACCCCAATGCTC 6780

Db 7061 GTGACCCGAGTATAGAGTGTAGTCCAAAGTGTGCTTGTGAGCTCTCTGACCCCAATGCTC 7120
Qy 6781 GAGGAAAGAGGCGGACCTTGAAGCTTGGATTCATTCAGAAATACATGCTTCCCAAGAGAGC 6840
Db 7121 GAGGAAAGAGGCGGACCTTGAAGCTTGGATTCATTCAGAAATACATGCTTCCCAAGAGAGC 7180
Qy 6841 TTCCCAACAGCTTTTACCGGCTGGGACAGGCTGATTAACAACCAACGCTTGTGGAATG 6900
Db 7181 TTCCCAACAGCTTTTACCGGCTGGGACAGGCTGATTAACAACCAACGCTTGTGGAATG 7240
Qy 6901 TGGAAAAGGCGAGATTAACAACGCGGCACTTGGCGGCTGTGCTCTCCCTCTCTAG 6960
Db 7241 TGGAAAAGGCGAGATTAACAACGCGGCACTTGGCGGCTGTGCTCTCCCTCTCTAG 7300
Qy 6961 AAAAAGGCGGCTTCCCAAGAGAGGCGGCAAGTGGGCTTGAATGAGACTCCATA 7020
Db 7301 AAAAAGGCGGCTTCCCAAGAGAGGCGGCAAGTGGGCTTGAATGAGACTCCATA 7360
Qy 7021 GAGATGCTCTTCAACAGTGGGCAATTAAGTCTTGGCCAGGCGGCGGCGGCGAT 7080
Db 7361 GAGATGCTCTTCAACAGTGGGCAATTAAGTCTTGGCCAGGCGGCGGCGGCGAT 7420
Qy 7081 TCAGGCTTTCACAGGAGGCGGCGGCGGATTCGCGAGTCAAGCGCTCTCTGATGAG 7140
Db 7421 TCAGGCTTTCACAGGAGGCGGCGGCGGATTCGCGAGTCAAGCGCTCTCTGATGAG 7480
Qy 7141 TTGGCCCTTTCCGAGAGAGGTTTCATCTCTTCAATGCCCCCTCCAGGGGAGCTTGA 7200
Db 7481 TTGGCCCTTTCCGAGAGAGGTTTCATCTCTTCAATGCCCCCTCCAGGGGAGGCTTGA 7540
Qy 7201 GATCAGAGCTTGAAGCTGAGAGTGAAGCTCCCAACCGGCGGCGGCGGCGGAGTGA 7260
Db 7541 GATCAGAGCTTGAAGCTGAGAGTGAAGCTCCCAACCGGCGGCGGCGGAGTGA 7600
Qy 7261 GCTCCGCGCTCGGACTCGGGGCTCTGATCTTGTCTCGAGAGAGAGCACTCGCTG 7320
Db 7601 ACCCGCGGCTCAGGCTCGGGGCTCTGATCTTGTCTCGAGAGAGAGCACTCGCTG 7660
Qy 7321 TGTGCTCCATGTCATATCTCTTGAACCGGGGCTTAAATCTCTTGTATCTCCGAAGAG 7380
Db 7661 TGTGCTCCATGTCATATCTCTTGAACCGGGGCTTAAATCTCTTGTATCTCCGAAGAG 7720
Qy 7381 GAGAAGTTACCGATTAACCTTGAAGCACTCCCTGTCGATATCAACAAGAGTGTAC 7440
Db 7721 GAGAAGTTACCGATTAACCTTGAAGCACTCCCTGTCGATATCAACAAGAGTGTAC 7780
Qy 7441 TGTACCAACAACAAGGCGCTTCAATAGGGCTTAAAGGTAACTTTGATAGATGCA 7500
Db 7781 TGTACCAACAACAAGGCGCTTCAATAGGGCTTAAAGGTAACTTTGATAGATGCA 7840
Qy 7501 GTGCTGAGCTCTTACATGAGCTCAATTAAGAGACATTAAGCTAGGCTTCAAGGTC 7560
Db 7841 GTGCTGAGCTCTTACATGAGCTCAATTAAGAGACATTAAGCTAGGCTTCAAGGTC 7900
Qy 7561 ACCGAAAGGCTCTCAACATGAGAGAGGCTTGCAGTTAAACCCCACTTGCAGAGA 7620
Db 7901 ACCGAAAGGCTCTCAACATGAGAGAGGCTTGCAGTTAAACCCCACTTGCAGAGA 7960
Qy 7621 TCTAAATATGAGGCTTGGGCGTAAAGAGTCCGAGCTTGTCCGGAGGGCGCTTAAAC 7680
Db 7961 TCTAAATATGAGGCTTGGGCGTAAAGAGTCCGAGCTTGTCCGGAGGGCGCTTAAAC 8020
Qy 7681 ATCAAGTCCGTGTGGAAGAGCTCTCTGAGAGATTCAGAAACATCAATTCACACAT 7740
Db 8021 ATCAAGTCCGTGTGGAAGAGCTCTCTGAGAGATTCAGAAACATCAATTCACACAT 8080
Qy 7741 ATGGCAAAAATGAGGCTTCTGCGTGAAGAGGCGGCGGCGGCGGCGGCGGCGGCGG 7800
Db 8081 ATGGCAAAAATGAGGCTTCTGCGTGAAGAGGCGGCGGCGGCGGCGGCGGCGGCGG 8140
Qy 7801 CTATGCTTAACTCTGAGCTCGGCTGAGGCTGCGGAGAGAGGCGGCTTATGACAT 7860
Db 8141 CTATGCTTAACTCTGAGCTCGGCTGAGGCTGCGGAGAGAGGCGGCTTATGATATC 8200

OY	7861	ACACAAAAA	CTTCCTCAGCGGTGATGGGGGCTTTATGATTCAGATATTC	CCCGCT	7920
Db	8201	ACACAAAAG	CTTCCTCAGCGGTGATGGGGGCTTTATGATTCAGATATTC	CCCGCT	8260
OY	7921	CAGCGGTAGAG	TTCTCTTGAAAGCATGGCGGAAAAAGAACCCATATGGGTTTTTCG	7980	
Db	8261	CAGCGGTGAG	TTCTCTTGAAAGCATGGCGGAAAAAGAACCCATATGGGTTTTTCG	8320	
OY	7981	TATGATACCCG	ATGCTTTGAGCTCAACGGTCACTGAGAGACATCAGAGACTGAGAGATCC	8040	
Db	8321	TATGATACCCG	ATGCTTTGAGCTCAACGGTCACTGAGAGACATCAGAGACTGAGAGATCC	8380	
OY	8041	ATATATCGGG	CTGCTGCTTCGTCGCGGAGAGGCCCACTGTCATACACTGCTAACTAG	8100	
Db	8381	ATATATCGGG	CTGCTGCTTCGTCGCGGAGAGGCCCACTGTCATACACTGCTAACTAG	8440	
OY	8101	AGACTTTAC	GTGGAGGGGCTATATGTTCAACAGCAAGGGCCAAACTCTGGGGTACAGGCT	8160	
Db	8441	AGACTTTAC	GTGGAGGGGCTATATGTTCAACAGCAAGGGCCAAACTCTGGGGTACAGGCT	8500	
OY	8161	TGCCCGCG	CAGCGGGGTGCTCAACCTATGAGCAACATACATGCTACGTGAAA	8220	
Db	8501	TGCCCGCG	CAGCGGGGTGCTCAACCTATGAGCAACATACATGCTACGTGAAA	8560	
OY	8221	GCTTTAGCG	CTTGTAAGCTGACAGAGATTAATCGGCCCAAGATGCTGTATGCGCGAT	8280	
Db	8561	GCTTTAGCG	CTTGTAAGCTGACAGAGATTAATCGGCCCAAGATGCTGTATGCGCGAT	8620	
OY	8281	GACTTGATG	TCATCTCAAGAAACCGAGGGGACCGAGAGAGACGAGCGGAACCTAGAGCC	8340	
Db	8621	GACTTGATG	TCATCTCAAGAAACCGAGGGGACCGAGAGAGACGAGCGGAACCTAGAGCC	8680	
OY	8341	TTACAGAG	GAGCTATGACAGGATATTCGCCCCCTCTGTGTGACCCCCCAGACCGGAGAT	8400	
Db	8681	TTACAGAG	GAGCTATGACAGGATATTCGCCCCCTCTGTGTGACCCCCCAGACCGGAGAT	8740	
OY	8401	GATCTGAG	CTGATTAACATCTTGCTCTCAAAATGTGTCTGTGGGCTTGGGCCACAGGC	8460	
Db	8741	GACCTGAG	CTGATTAACATCTTGCTCTCAAAATGTGTCTGTGGGCTTGGGCCACAGGC	8800	
OY	8461	CGCGCAG	ATACTCTGACAGACCTTACCACTCCAAATGCGCCGGGCTGCGTGGGA	8520	
Db	8801	CGCGCAG	ATACTCTGACAGACCTTACCACTCCAAATGCGCCGGGCTGCGTGGGA	8860	
OY	8521	ACAGTTAG	ACACTCCCTGTCAATTCATGCTGGGAAACATCATCAGTACGCCCCAGC	8580	
Db	8861	ACAGTTAG	ACACTCCCTGTCAATTCATGCTGGGAAACATCATCAGTACGCCCCAGC	8920	
OY	8581	ATATGGG	CTCGATGCTGTATGACATCACTTTCTCTCATTTGATGCTCAAGACAGC	8640	
Db	8921	ATATGGG	CTCGATGCTGTATGACATCACTTTCTCTCATTTGATGCTCAAGACAGC	8980	
OY	8641	CTGACAG	AACTCAACTTTGAGATGTACGAGGCGGTACTCCGTAAGTCCCTTGAC	8700	
Db	8981	CTGACAG	AACTCAACTTTGAGATGTACGAGGCGGTACTCCGTAAGTCCCTTGAC	9040	
OY	8701	CTCCAGCT	ATATGAAAGGTTACATGGGCTTGACGCTTTTCTCTGCAACATACACT	8760	
Db	9041	CTCCAGCT	ATATGAAAGGTTACATGGGCTTGACGCTTTTCTCTGCAACATACACT	9100	
OY	8761	CCCGCAG	AATCTACACGCGGTGCTCAGGCCCTCAGAAAACTTGGGGGCGCACCCCTCGA	8820	
Db	9101	CCCGCAG	AATCTACACGCGGTGCTCAGGCCCTCAGAAAACTTGGGGGCGCACCCCTCGA	9160	
OY	8821	GGGTGAA	AGCGCGGACATGACATGACATGACGCGGCTCCATCTCCGTTGGGGGAGAGCG	8880	
Db	9161	GGGTGAA	AGCGCGGACATGACATGACATGACGCGGCTCCATCTCCGTTGGGGGAGAGCG	9220	
OY	8881	GGCGTTTG	CGGCTGATCTCTTCAATTTGGGGGTGGAAGCCAGACTCAAACTCACTCA	8940	
Db	9221	GGCGTTTG	CGGCTGATCTCTTCAATTTGGGGGTGGAAGCCAGACTCAAACTCACTCA	9280	

Qy	8941	TTGCGGAAGGCGCCCTTCGATTTATTCAGCTGATTACCGTGGCGCGCGCGGAGGCG	9000
Db	9281	TTGCGGAAGGCGCCCTTCGATTTATTCAGCTGATTACCGTGGCGCGCGCGGAGGCG	9340
Qy	9001	GACATTATACACAGCGTGC CGCGCCGACCCCGCTATTATGCTCTTTGGCGTAACTCTA	9060
Db	9341	GACATTATACACAGCGTGC CGCGCCGACCCCGCTATTATGCTCTTTGGCGTAACTCTA	9400
Qy	9061	CTTTTGTAGGGGTAGGCTTTTCTTACTCCCGGCTGG	9099
Db	9401	CTTTTGTAGGGGTAGGCTTTTCTTACTCCCGGCTGG	9439

RESULT	6
E07361	
LOCUS	E07361
DEFINITION	gRNA of Hepatitis non-A non-B virus, HCV-J6.
ACCESSION	E07361
VERSION	E07361.1 GI:5708558
KEYWORDS	JP 1994121689-A/1.
SOURCE	unidentified
ORGANISM	unidentified
REFERENCE	unclassified.
AUTHORS	1 (bases 1 to 9589)
TITLE	Okamoto,H. and Nakamura,T.
JOURNAL	NON-A NON-B HEPATITIC VIRUS GENE, POLYNUCLEOTIDE, ANTIGEN AND ANTIBODY DETECTION SYSTEM, Patent: JP 1994121689-A 1 06-MAY-1994;
PAT	29-SEP-1997

COMMENT	OS	Hepatitis non-A non-B virus
	PN	JP 1994121689-A/1

PI OKAMOTO HIROAKI, NAKAMURA TETSUO
PC C12N15/51, A61K39/29, A61K39/395, C07K31/00, C12P21/02,
PC C12P21/08,
PC G01N33/53, G01N33/576, G01N33/577//A61B10/00;

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

FT	FT
/organism='Hepatitis non-A non-B virus'	FT

FEATURES	Location/Qualifiers
source	1. .9589

Query Match 93.4%; Score 8494.2; DB 6; Length 9569;

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523</
--	---	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------

Db 341 ATGAGCACAAATCCTAAACCTCAAGAAAAACCAAGAACAACCAACCGTCGCCACAA 400

121 GGGCCAGTTGGTGTGCGCGACACAGGAAGACTTCGAGCGGTCCACGCCACGTGA 180

QUESTION

Db 521 AGCGCCAGCCCATCCCTAAGGATCGGCGCTCCACTGGCAATCCTGGGAAACCAGGA 580

OY	241	TACCCTTGAGCCCTTAATAGGGAAATAGAGGACTTCGGCTGGGAGAGATGGCTCTCTGTCCCC	300
Db	581	TACCCCTG3CCCCCTATATACGGAAATAGAGGAACTTCGGCTGGGAGAGATGGCTCTCTGTCCCC	640
OY	301	CGAGGTTCCCGGCTCTTTGGGGGCCCCCAATGACCCCCGGGCAATGAGTGCAGCAAGTGGAGT	360
Db	641	CGAGGTTCCCGTCCCTCTTTGGGGGCCCCCAATGACCCCCGGGCAATGAGTGCAGCAAGTGGAGT	700
OY	361	AAGGTCATCGATACCTTAACGTGCGGCTTTGGCCGACCTCATGGGGTACATCCCTGTGCTG	420
Db	701	AAGGTCATCGATACCTTAACGTGCGGCTTTGGCCGACCTCATGGGGTACATCCCTGTGCTG	760
OY	421	GGCGCCCCCGCTCGGGGGGGTGGCCAGACTCTCGGCAATGGCGTGAAGTCTTGAAGAC	480
Db	761	GGCGCCCCCGCTCGGGGGGGTGGCCAGACTCTCGGCAATGGCGTGAAGTCTTGAAGAC	820
OY	481	GGGGTTAATTTTGGCAACAGGGAACTTAACCGGTTGCTCTTTTCTATCTTCTGTCGGCC	540
Db	821	GGGGTTAATTTTGGCAACAGGGAACTTAACCGGTTGCTCTTTTCTATCTTCTGTCGGCC	880
OY	541	CTGCTGTCCTGCATCAACCAACCCGGTCTCGCTGCGCAAGTGAAGAACATCAAGTACCGGC	600
Db	881	CTGCTGTCCTGCATCAACCAACCCGGTCTCGCTGCGCAAGTGAAGAACATCAAGTACCGGC	940
OY	601	TACATGTGATCAACAGACTGCAACCAATGACATTAATCTTG3AGCTTCAGGCTGCTGTC	660
Db	941	TACATGTGATCAACAGACTGCAACCAATGATAGCAATTAATCTTG3CAATTCAGGCTGCTGTC	1000
OY	661	CTCAACGTCCTCCGGGTGCTCCCGGGGAGAAAGTGGGGAATGCAATCAATGCTGAGTA	720
Db	1001	CTCAACGTCCTCCGGGTGCTCCCGGGGAGAAAGTGGGGAATCAATCAATGCTGAGTA	1060
OY	721	CCGGTCTTCAACGGAATGTGGCGGTGCAAGCGGCCCGGCGCTCTACGCAAGGCTTGCAGACG	780
Db	1061	CCGGTCTTCAACGGAATGTGGCGGTGCAAGCGGCCCGGCGCTCTACGCAAGGCTTGAAGACG	1120
OY	781	CACATCGACATGTTGTGATGTCCGCCACGCTGCTCTGCTGCTCTGATGTTGGGAGACTCTC	840
Db	1121	CACATTTGACATGTTGTGATGTCCGCCACGCTGCTGCTGCTCTTCTTACGTGGGGAACCTC	1180
OY	841	TGCGGTGGGGTGTGATGCTCGCAGGCCCAATGTTTCATTTGCTCGCGCAGACACCATGTGTT	900
Db	1181	TGCGGTGGGGTGTGATGCTCGCAGGCCCAATGTTTCATTTGCTCGCAGACACCATGTGTT	1240
OY	901	GTCGAAGCTGCAATTTGCTCCATCTCAACCTGTAACATCACTGGACACCGCATGGCATGG	960
Db	1241	GTCGAAGCTGCAATTTGCTCCATCTCAACCTGTAACATCACTGGACACCGCATGGCGTGG	1300
OY	961	GACATGATGATGAATCGTGTGCGCCACCGGCTACATGATCTTGGGGTACGCGATGCGTGTCT	1020
Db	1301	GACATGATGATGAATCGTGTGCGCCACCGGCTACATGATCTTGGGGTACGCGATGCGGCTCT	1360
OY	1021	CCCCAGGTCATTTATGACATCAATTAAGCGGGGCTCATTTGGGGGCTCATGTTGGCTTGCC	1080
Db	1361	CCCCAGGTCATTTATGACATCAATTAAGCGGGGCTCATTTGGGGGCTCATGTTGGCTTGACC	1420
OY	1081	TACTTCTCTATAGCAGGAGCGTGGGCGAAAGTGTGTGCATCTTCTGTTGGCGCGCGGG	1140
Db	1421	TACTTCTCTATAGCAGGAGCGTGGGCGAAAGTGTGTGCATCTTCTGTTGGCGCGCGGG	1480
OY	1141	GTCGACGCGGCAACCCATATCTTTGGGGGTTCTGCGCGCAGACCAACCGGCGGCTCAAC	1200
Db	1481	GTCGACGCGGCAACCCATATCGTTGGGGGTTCTACCGCGGCAATACCGCAGGACCTCAAC	1540
OY	1201	AGCTTATTTGACATGGGCCCCCAGGCAAAAATCCAGCTCATCAACCAACCAATGGCAGTGG	1260
Db	1541	AGCTTATTTGACATGGGCCCCCAGGCAAAAATCCAGCTCATCAACCAACCAATGGCAGTGG	1600
OY	1261	CACATCAACCGCACCGCCTGTAACTGCAATAGACTCTTGGACACCGGCTTATGCGGCTCT	1320
Db	1601	CACATCAACCGCACCGCCTGTAACTGCAATAGACTCTTGGACACCGGCTTCTGCGCTCA	1660
OY	1321	CTGTTTCAACACCAAGCTTCAACTGTCAGAGATGTCCGAAACGATGTCCGCTGCGCTC	1380

Db	1661	CTGTTCTACACCCACAGCTTCACTGCTCAGAGATGTCACCAAGCCATGTCGGCTCGCG	1720
Qy	1381	AGTATCGAGGCGCTTCCGGGTGGGATGCGGGCGCTTGGCAATATGAGATTAATGTCACCAAT	1440
Db	1721	AGTATCGAGGCGCTTCCGGGTGGGATGCGGGCGCTTACAAATATGAGCAATATGTCACCAAT	1780
Qy	1441	CCAGAGATATAGAACCTCTATTGCTGCACTAACCCACCAAGCAGTGTGGCGTGTCTCC	1500
Db	1781	CCAGAGATATAGAACCGTATTTGCTGGCACTTACCCACCAAGCAGTGTGGGTATGTCCTCC	1840
Qy	1501	GCGAAGACTGTGTGTGGGCCAGTGTACTGTATTACCCCGAGCCGACGTATGATGGGACG	1560
Db	1841	GCGAGCTTGTGTGTGGGCCAGTGTACTGTATTACCCCGAGCCGACGTATGATGGGATCG	1900
Qy	1561	ACCGACAGGCTTGGAGGCGCCACTTACACGTGGGGGAGATGAGACAGATGTCTTCTTA	1620
Db	1901	ACCGATAGACTTGGAGGCGCCACTTACACGTGGGGGAGATGAGACAGATGTCTTCTTA	1960
Qy	1621	TTGAAACAGACTCGACCAACCGCTGGGGGTCAATGTTGGCTGACAGTGAATGAACCTTCTT	1680
Db	1961	TTGAAACAGACTCGACCAACCGAGGGGTCAATGTTGGCTGACAGTGAATGAACCTCACT	2020
Qy	1681	GGCTACACCAAGACTTGGCGGCGCACACCCCTGCTTACTAGACTGACTTCAACGCCACG	1740
Db	2021	GGCTACACCAAGACTTGGCGGCGCACACCCCTGCGCATTAAGACTGACTTCAATGCGACG	2080
Qy	1741	ACGGACCTGTTGTGCTCCCGACGAGACTGTTTTAGGAAGATCTGTATACACTTACCTCTCAA	1800
Db	2081	ATGACACTTGTGTGTGCTCCCGACGAGACTGTTTTAGGAAGATCTCTGTATACCACTACATCAA	2140
Qy	1801	TGCGGCTCTGCGGCCCTTGGCTCAGCGCAAGGTGCTGTATGACTAACCCCTACAGGCTCTGG	1860
Db	2141	TGTGGCTCTGCGGCCCTTGGCTCAGCGCAAGGTGCTGTATGACTAACCCCTACAGGCTCTGG	2200
Qy	1861	CATTACCCCTGACACAGTTAACTATACATCTTCAAAATAAGATGTTATGGGAGGGGTT	1920
Db	2201	CATTACCCCTGACACAGTTAACTATACATCTTCAAAATAAGATGTTATGGGAGGGGTT	2260
Qy	1921	GAGACACGGCTCAACGGCGCATGCAATTTCACTCGTGGGATGGTTGCAACTTGGAGAGAC	1980
Db	2261	GAGACACGGCTCAACGGCGCGTGGTGAATTTCACTCGTGGGATGGTTGCAACTTGGAGAGAC	2320
Qy	1981	AGAGACAGAAATCAACTGTCTCTCTTTGTTGCACTCCACACGGAATGGGCAATTTTAACT	2040
Db	2321	AGAGACAGAAATCAACTGTCTCTCTTTGTTGCACTCCACACGGAATGGGCAATTTTAACT	2380
Qy	2041	TGCTCTTACTCGACACTGCGCGCTTGTGCACTGTCTTCTCAACTCAACCAAAACATC	2100
Db	2381	TGCACTTACTCGACACTGCGCGCTTGTGCACTGTCTTCTTCAACCTCAACCAAAACATC	2440
Qy	2101	GTGACGCTAACATTCAATGTAATGGCTATACACGTGCCCTCAAAAAATCAATGTCGCAATG	2160
Db	2441	GTGACGCTAACATTCAATGTAATGGCTATACACGTGCTTCAAAAAATCAATGTCGCAATG	2500
Qy	2161	GAGTGGGTAAATACCTTATTCCTGCTCTTATGCGGACGCGCAGGGTTTGCGCTGCTTAATG	2220
Db	2501	GAGTGGGTAAATACCTTATTCCTGCTCTTATGCGGACGCGCAGGGTTTGCGCTGCTTAATG	2560
Qy	2221	ATGCTCATCTTGTGTTGGGCTCAGGCCGAGACAGCACTAGAGAAAGCTGTCAATCTTGGACGCT	2280
Db	2561	ATGCTCATCTTGTGTTGGGCTCAGGCCGAGACAGCACTAGAGAAAGCTGTCAATCTTGGACGCT	2620
Qy	2281	GCGAGGCCAGATGACCTGCAATGGCTTCTATATTTGTTGATCTTTTTCGAGGCGCTGTGG	2340
Db	2621	GCGAGGCCAGATGACCTGCAATGGCTTCTATATTTGTTGATCTTTTTCGAGGCGCTGTGG	2680
Qy	2341	TACATCAAGGGTCGGGTAGTCCCTTATAGTAACTATTCCTCACTGGCTGTGTGCTCTT	2400
Db	2681	TACATCAAGGGTCGGGTAGTCCCTTATAGTAACTATTCCTCACTGGCTGTGTGCTCTT	2740
Qy	2401	AGCTTACTGCTCTTAGCAATTGCCCAACAGGCTTATGCTTATGACGATCTGTGCAATGCG	2460

Db	2741	ggccctactgctcttagactttggccccaagagccttatgctatgacgcattctgtacatgct	2800
OY	2461	cagatagagagcgccctctctgctgtaataatcactctctttactctcaccgccggtatpaa	2520
Db	2801	cagatagagagagcctctgttggtatctgatactctcttttatacctcaccgccggtatpaa	2860
OY	2521	accctctcagcccggttttttgngtggcttgatctcttgcagcccttgagggaagctatg	2580
Db	2861	accctctcagcccggtttctgtgggtggcttgatctcttgcagcccttgagggaagctatg	2920
OY	2581	gtccagagatgggacacacctatgacaggtggcgcgctgattgacatattgggccc	2640
Db	2921	gtccagagatgggacacacctatgacaggtggcgcgctgattgacatattgggccc	2980
OY	2641	gtccgcatattcttaccagagtggtgctgtttgcataatpccaaatggagctcttgagcgagctt	2700
Db	2981	gtccgcatattcttcccgcggtgtggtgttttgacatpaccaaatggagctcttgagcgagctt	3040
OY	2701	gggcctgcttactctcttataaagtgctttgacgcgcgctgaccgtatcttgcagagcgctac	2760
Db	3041	gggcctgcttactctcttataaagtgctttgacgcgcgctgaccgtatcttgcagagcgctac	3100
OY	2761	gctctactgagagatgtgacacatggcagagatcttcggcggggggagaggtatagtcacatg	2820
Db	3101	gctctactgagagatgtgacacatggatpaaaggacatctgcgggggggagaggtatagtcacatg	3160
OY	2821	gcgctactagcccttggcagagtgagctgagcattacattatgacacactcaccctatg	2880
Db	3161	gngctactagcccttggcagagtgagctgagcattacattatgacacactcaccctatg	3220
OY	2881	tcggattggcgctgctagtgagccctgcggagcctcgcgcgctgcgcgttagcctatcattcttc	2940
Db	3221	tcggattggcgctgctagtgagccctgcggagcctcgcgcgctgcgcgttagcctatcattcttc	3280
OY	2941	agtcgcatgagaaamaaagtcattgtctggggagcgagagacagctgcttgaggggacatt	3000
Db	3281	agtcgcatgagaaamaaagtcattgtctggggagcgagagacagctgcttgaggggatgac	3340
OY	3001	tttaccggaacttccggtgtccgcgcgcgacttgatgggaaggtcttcttgagccacgctgat	3060
Db	3341	tttaccggaacttccggtgtccgcgcgcgacttgatgggaaggtcttcttgagccacgctgat	3400
OY	3061	ggccttactctccaaaggggtgagatcttctgcgcccatcactgctttagccccagagaga	3120
Db	3401	ggccttactctccaaaggggtgagatcttctgcgcccatcactgctttagccccagagaga	3460
OY	3121	cgtagcccttattgggacacattagtgctgagacatgacggggcgcgacamaacagaaacagcct	3180
Db	3461	cgtagcccttattgggacacattagtgctgagacatgacggggcgcgacamaacagaaacagcct	3520
OY	3181	ggggaaatttcaggtcctgtgccacagctcactcagctcttcttcggaaactccatctcgggg	3240
Db	3521	ggggaaatttcaggtcctgtgccacagctcactcagctcttcttcggaaactccatctcgggg	3580
OY	3241	gttttgtagaactgtctaacatggaactgagcaaacagactctgagccgggtccg	3300
Db	3581	gttttgtagaactgtctaacatggaactgagcaaacagactctttagccgggtccg	3640
OY	3301	gtcagcagatgtaactccagtgcttagagggggaatttagtaggtggccacagcccccctggg	3360
Db	3641	gtcagcagatgtaactccagtgcttagagggggaatttagtaggtggccacagcccccctggg	3700
OY	3361	actaatcttttgagacccgtgacacgtgtggaaggcggtgcacctgtacttggtcaccgggaac	3420
Db	3701	actaatcttttgagacccgtgacacgtgtggaaggcggtgcacctgtacttggtcaccgggaac	3760
OY	3421	gctgtagtcatcccgactcgaaagacggcgggacaaacggggagcgctactctcccgaga	3480
Db	3761	gctgtagtcatcccgactcgaaagacggcgggacaaacggagagcgctactctcccgaga	3820
OY	3481	cctcttttgcactttgaaagggtctctcagagagcccggtgtctatgtgccacaggggccaagct	3540
Db	3821	cctcttttgcactttgaaagggtctctcagagagcccggtgtctctgcccacagggcacaagct	3880

QY	3541	GTCCGAGCTCTTCGGGGAGAGCTGTGTCTCTTCGGGGCGTGGCTAAGTCGATAGATTTC	3600
Db	3881	GTCCGGGCTCTTCGGGGAGCGCGTGTGTCTCCCGGGGCGTGGCCAAAGTCATAGATTTC	3940
QY	3601	CCCGTTGAGCACTCGACATCGTCAGCCGGTCCCCACCTTATAGTGACAAAGCACACCA	3660
Db	3941	CCGGTTAGACACTTGACATGTGTCACTCGGTCCCGACCTTTAGTGACAAAGCACACCA	4000
QY	3661	CCGTGCTGTGCCAGACCTTATCAGTGGGGTACTTGATGATGCCCACTGGCAGTGGAAAG	3720
Db	4001	CTGTCTGTGCCCAAACTTATCAGGTGGGACTTATATCCCGACTGGTATGGGAAAG	4060
QY	3721	AGACCAAAAGTCCGTGCGATATAGCTGCTCAGGGGTATAAAGTCTAGTCTTAATCCC	3780
Db	4061	AGACCAAAAGTCCGTGCGGTATGCGGCTCAGGGGTATCAAAAGTCTAGTCTTAATCCC	4120
QY	3781	TCAGTGGCTGCACCTCTGGGGTTTGGGGCGTACTTGTCTTAAAGGCATATGGCATATCCC	3840
Db	4121	TCGTGTGCTGTCACCTCGGGGTTTGGGGCGTACTTGTCTCAAGGCACTATGGCATCAATCCC	4180
QY	3841	AACATTAGGACTGGAGTCAAGGACTGTACGACCGGGGCGCCATCAGTACTCCACATAT	3900
Db	4181	AACATTAGGACTGGGGTTCAGGACTGTACGACCGGGGCGCCATCAGTACTCCACATAT	4240
QY	3901	GGCAAAATTCCTCGCGATGGGGGCTGTGCGGGCGGCCCTTACGACATCATATATGTGAT	3960
Db	4241	GGCAAAATTCCTCGCGATGGGGGCTGTGCGAGGGCGGCTTATGACATCATATATGTGAT	4300
QY	3961	GAATGCCATGCGGTGACTCTACCAACATCTTGGACTGGMAACAGTCTTGTATCAAGCA	4020
Db	4301	GAATGCCATGCGGGGACTCTACCAACATTCCTGGCATGGMAACAGTCTCTGTATCAAGCA	4360
QY	4021	GAGACAGTGGGGTCAACTAATCTGTCTGGCTTACACTTACGCCCCCTGGGTCAGTACA	4080
Db	4361	GAGACAGCGGGGTCAAGCTTAACTGTACTGGCTTACGCTTACGCCCCCGGGTCAGTACA	4420
QY	4081	ACCCCCCAACCCCAATAGAGAGAGGTGGCCCTTGGGCAAGAGGGCGAGATCCCTTCTAT	4140
Db	4421	ACCCCCCAACCCCAATAGAGAGAGGTGGCCCTCGGGAGAGAGGTGAGATCCCTTCTAT	4480
QY	4141	GGGAGGGCGAATCCCTGTCTTATCATAGAGGAAGAAAGACATCGATCTTCTGCACTTCA	4200
Db	4481	GGGAGGGCGAATCCCTGTCTCATATCATAGGAGAAAGAACATCTTATCTTCTGCACTCA	4540
QY	4201	AAGAAAAAGTGTACGAGCTCGCGCGGCGCCCTTGCGGGTATGGGCTTGAATCTAGTGGCA	4260
Db	4541	AAGAAAAAGTGTACGAGCTCGCGCGGCGCCCTTGCGGGTATGGGCTTGAACGCGAGTGGCA	4600
QY	4261	TACTTACAGAGGTTGAGAGCTCTCCGTAAATACCACTCAGGGAGACGTATGTCTGTGGCC	4320
Db	4601	TACTTACAGAGGGCTGAGAGCTCTCCGTAAATACCACTCAGGGAGACGTATGTCTGTGCC	4660
QY	4321	ACCGACGCCCTCATGACAGGGTATACTGAGGGACTTTGACTCCGATTCGACTCGACAAAGTA	4380
Db	4661	ACCGACGCCCTCATGACAGGGGTTTACTGAGAGACTTTGACTCCGATTCGACTCGAAAGTA	4720
QY	4381	GCGGTCACTCAAGTTGTAGACTTCAATTAGACCCCAATTCACCATTAACCAACAGATT	4440
Db	4721	GCGGTCACTCAAGTTGTAGACTTCAAGCTTGGACCCCAATTCACCATTAACCAACAGACT	4780
QY	4441	GTCCCTCAAGACGTGTCTCAAGTACGAGGCGCGGGGTTGSCACGGGTATAGGGGAAAGCTG	4500
Db	4781	GTCCCTCAAGACGTGTCTCAAGTACGAGGCGCGGGGCGGCAAGGCGAGGGGAAAGACTG	4840
QY	4501	GGCATTTTATAGTATGTTTCCACTGTGTAGCGAGCCTCAGGAATGTTTGAACAGTGTAGTG	4560
Db	4841	GGTATTTTATAGTATGTTTCCACTGTGTAGAGAGCCTCAAGGAATGTTTGAACAGTGTAGTG	4900
QY	4561	CTCTGTAGTGTCTTACGACGAGGGGCGCGATGTATATAGCTTCAACCATCGGAAACCAAC	4620
Db	4901	CTCTGTAGTGTCTTACGATGAGGGGCGCGATGTATATAGCTTCAACCAAGCGGAAACCAAC	4960

QY 4621 GTCAGGCTGAGGCGGTATTTCAACAGCCCGGTTGCTGTGCCAAGACATCTTGAG 4680
|||||
Db 4961 GTCAGGCTGAGGACATATTTCAACACACCGTTTGTCTGTGCCAAGACATCTTGAG 5020
|||||
QY 4681 TTTTGGAGGCGAGTTTTCACCGGCTTCAACACATAGATGCCCATTTCTTTCCAAACA 4740
|||||
Db 5021 TTTTGGAGGCGAGTTTTCACCGGCTTCAACACATAGATGCCCATTTCTTTCCAAACA 5080
|||||
QY 4741 AAGCAATCGGGGGGAAATTTGCGATCTTAAACAGCCCTACAGGCTTACAGTGTCCGTAG 4800
|||||
Db 5081 AAGCAATCGGGGGGAAATTTGCGATCTTAAACAGCCCTACAGGCTTACAGTGTCCGTAG 5140
|||||
QY 4801 GCCAAAGCCCCCCCCCGTCTCTGGAGCATGATGTGAAGTGTGACTCGACTCAAGCCC 4860
|||||
Db 5141 GCCAAAGCCCCCCCCCGTCTCTGGAGCATGATGTGAAGTGTGACTCGACTCAAGCCC 5200
|||||
QY 4861 ACACTGTGGGCCCCCAACCTCTCTGTACCGGTTGGGCTTGTAAACAAGAGTACCC 4920
|||||
Db 5201 ACACTGTGGGCCCCCAACCTCTCTGTACCGGTTGGGCTTGTAAACAAGAGTACCC 5260
|||||
QY 4921 CTCACACATCCCGTGAAGAAATACATCGCACCTTGCATGACAGCCGCTTGAAGTCAAG 4980
|||||
Db 5261 CTCACACATCCCGTGAAGAAATACATCGCACCTTGCATGACAGCCGCTTGAAGTCAAG 5320
|||||
QY 4981 ACCAGACATGGGCTTGTGGCAGGGGAGTCTTGGCGGCGCTGCGCGGATTTGCTGGCG 5040
|||||
Db 5321 ACCAGACATGGGCTTGTGGCAGGGGAGTCTTGGCGGCGCTGCGCGGATTTGCTGGCG 5380
|||||
QY 5041 ACCGSGTGTGTTCATCATCGGCGGCTTGCACATTTAACAGGAGCGGCTGTGGCGCG 5100
|||||
Db 5381 ACCGSGTGTGTTCATCATCGGCGGCTTGCACATTTAACAGGAGCGGCTGTGGCACG 5440
|||||
QY 5101 GACAAAGAGTCTCTATATAGGCTTTTGAATGAGATGGAAGAAATGTCTTAGGGGCT 5160
|||||
Db 5441 GAAAGAGAGTCTCTATATAGGCTTTTGAATGAGATGGAAGAAATGTCTTAAAGGGCT 5500
|||||
QY 5161 CTCATTGAAGAGGGGAGCGGATAGCCGAGATCTGAAGTCCAAAGTCCAAAGCTTAATG 5220
|||||
Db 5501 CTCATTGAAGAGGGGAGCGGATAGCCGAGATCTGAAGTCCAAAGTCCAAAGCTTAATG 5560
|||||
QY 5221 CAGCAAGCTTCCAAACAGCTCAAGACATACAAACCACTGTGAGGCTTATGAGGCCAAG 5280
|||||
Db 5561 CAGCAAGCTTCCAAACAGCTCAAGACATACAAACCACTGTGAGGCTTATGAGGCCAAG 5620
|||||
QY 5281 GTTGAACATTTCTGGGCGCAACCATGTGGAATCTTATGCGGCACTCAATACCTTGCA 5340
|||||
Db 5621 GTTGAACATTTCTGGGCGCAACCATGTGGAATCTTATGCGGCACTCAATACCTTGCA 5680
|||||
QY 5341 GGAATATCAACACTGCGCAGGGAACCTGTGCTGATGCTTCCATGATGGCATTCAGTCCGCG 5400
|||||
Db 5681 GGAATATCAACACTGCGCAGGGAACCTGTGCTGATGCTTCCATGATGGCATTCAGTCCGCG 5740
|||||
QY 5401 CTGACAGTCCGCTGTCAACAGACCATATCTTCTCAACATTTTGGGGGCTGTGCTA 5460
|||||
Db 5741 CTGACAGTCCGCTGTCAACAGACCATATCTTCTCAACATTTTGGGGGCTGTGCTA 5800
|||||
QY 5461 GCATTCCTCAATTTGCAACACCCCGGGGGGCACTGGGCTTGTGTCAATGGCTTATGTGGA 5520
|||||
Db 5801 GCATTCCTCAATTTGCAACACCCCGGGGGGCACTGGGCTTGTGTCAATGGCTTATGTGGA 5860
|||||
QY 5521 GCTGCGGTAGGCGAGTATAGGCTTAGGTAGGTGCTAGTGAACATCTGAGCAGAGGTATGT 5580
|||||
Db 5861 GCTGCGGTAGGCGAGTATAGGCTTAGGTAGGTGCTAGTGAACATCTGAGCAGAGGTATGT 5920
|||||
QY 5581 GCGGGCATTTTGGGGGCTCTGTGCAATTCAGATCATGTCTGCGCAAGAGCCCTCCAGT 5640
|||||
Db 5921 GCGGGCATTTTGGGGGCTCTGTGCAATTCAGATCATGTCTGCGCAAGAGCCCTCCAGT 5980
|||||
QY 5641 GAGAGATGTCTCAACTTTGTGCTGCGCAATTTCTGCTCCGGGCTTGTGTGAATGTGGAATG 5700
|||||
Db 5981 GAGAGATGTCTCAACTTTGTGCTGCGCAATTTCTGCTCCGGGCTTGTGTGAATGTGGAATG 6040
|||||
QY 5701 ATCTGCGGCGCATTTGTGCGCGACATGTGGACCGGGGGAGAGCGCGCTTCAATGATG 5760
|||||

Db 6041 ATCTGCGGCGCATTTGTGCGCGACATGTGGACCGGGGGAGAGCGCTTCAATGATG 6100
|||||
QY 5761 AATAGACTATTTGCTTGTGCTTCCAGAGGAATACGTGCGCCCAACCACTAGTGTAG 5820
|||||
Db 6101 AATAGACTATTTGCTTGTGCTTCCAGAGGAATACGTGCGCCCAACCACTAGTGTAG 6160
|||||
QY 5821 GAGTGGATGCGGTGAGGAGTGTGACCAACTACTTGGGCTCCCTTACATTAACAGCTG 5880
|||||
Db 6161 GAGTGGATGCGGTGAGGAGTGTGACCAACTACTTGGGCTCCCTTACATTAACAGCTG 6220
|||||
QY 5881 CTCAGAAAGCTCCCAACTGTGATTAATCTGAGGACTGCCCCATCCATGCGGCGCTGTG 5940
|||||
Db 6221 CTCAGAGACTCCCAACTGTGATTAATCTGAGGACTGCCCCATCCATGCGGCGCTGTG 6280
|||||
QY 5941 CTCGCGATGTGTGGGACTGGGTTTGCACCATCTTAAACAGCTTTAAATTTGGCTGACC 6000
|||||
Db 6281 CTCGCGATGTGTGGGACTGGGTTTGCACCATCTTAAACAGCTTTAAATTTGGCTGACC 6340
|||||
QY 6001 TCCAAATTAATCCCAAGATGCCCCGCTCCCTTGTCTCCGTCAAAAAGGGGTACAG 6060
|||||
Db 6341 TCCAAATTAATCCCAAGATGCCCCGCTCCCTTGTCTCCGTCAAAAAGGGGTACAG 6400
|||||
QY 6061 GGGGTGTGGGCGGCGCACTGSCATCATGACCAACGSGTGTCTTGGCGGCAATATCTCT 6120
|||||
Db 6401 GGGGTGTGGGCGGCGCACTGSCATCATGACCAACGSGTGTCTTGGCGGCAATATCTCT 6460
|||||
QY 6121 GGGCAATGTCCGCTTGGGCTTCATGAGAAATCAAGGGGCTTAAAGCTTCAATATCTG 6180
|||||
Db 6461 GGGCAATGTCCGCTTGGGCTTCATGAGAAATCAAGGGGCTTAAAGCTTCAATATCTG 6520
|||||
QY 6181 CAGGGGACCTTTCTATCAATTTTTCACAGGAGGGGCAAGTGTGCGGAAACCGGCGCA 6240
|||||
Db 6521 CAGGGGACCTTTCTATCAATTTTTCACAGGAGGGGCAAGTGTGCGGAAACCGGCGCA 6580
|||||
QY 6241 AACTTTAAGTCTGSCATCTGAGGGTGTGGGCTTCAAGATCAAGGAGTGAAGCAGAC 6300
|||||
Db 6581 AACTTTAAGTCTGSCATCTGAGGGTGTGGGCTTCAAGATCAAGGAGTGAAGCAGAC 6640
|||||
QY 6301 GGGTCAATCACTACATTAACAGAGCTCAACCATGATTAATTTAAATGCTTCCGCACTA 6360
|||||
Db 6641 GGGTCAATCACTACATTAACAGAGCTCAACCATGATTAATTTAAATGCTTCCGCACTA 6700
|||||
QY 6361 CCTCTCCGAGTCTTTTCTGGGAGGAGCGAGTGCAGATCCATAGTGTGCCCCACA 6420
|||||
Db 6701 CCTCTCCGAGTCTTTTCTGGGAGGAGCGAGTGCAGATCCATAGTGTGCCCCACA 6760
|||||
QY 6421 CCGAAGCGTTTTTCCGGGATGAGGTCTGCTTCTGCTGGGCTTAAATTCATTTGTCTG 6480
|||||
Db 6761 CCGAAGCGTTTTTCCGGGATGAGGTCTGCTTCTGCTGGGCTTAAATTCATTTGTCTG 6820
|||||
QY 6481 GGGTCCGAGTCTTCTGCGACCTTGAACCCGACAGAGATTAATGATGTCTATGCTAACA 6540
|||||
Db 6821 GGGTCTTCACTCTCTTGCATCTGAACTGTACACAGAGTATTAAGTCTCAAGCTTAACA 6880
|||||
QY 6541 GATCATCTCATATCAAGGAGGAGCTGAGAGCGGCGGTTTATGCGGGGGGTACCCCA 6600
|||||
Db 6881 GATCATCTCATATCAAGGAGGAGCTGAGAGCGGCGGTTTATGCGGGGGGTACCCCA 6940
|||||
QY 6601 TCCGAGGCAAGCTCTTCCGAGAGCCAGCTATCGGCACTATGCTGCGAGCCTGACCC 6660
|||||
Db 6941 TCCGAGGCAAGCTCTTCCGAGAGCCAGCTATCGGCACTATGCTGCGAGCCTGACCC 7000
|||||
QY 6661 ACCCAAGGCAAGCTTATGATGTGACATGTGTGATGCTAACCCTGTTATGAGGGGCGAT 6720
|||||
Db 7001 ACCCAAGGCAAGCTTATGATGTGACATGTGTGATGCTAACCCTGTTATGAGGGGCGAT 7060
|||||
QY 6721 GTGACTCGGATAGAGCTGAGGTCCAAAGTGTGTCTGTGGAATCTCTCGACCCCAATGATG 6780
|||||
Db 7061 GTGACTCGGATAGAGCTGAGGTCCAAAGTGTGTCTGTGGAATCTCTCGACCCCAATGATG 7120
|||||
QY 6781 GAGAAAGAGAGCACTTGAAGCTTGTGATACATGCAATTAACATGCTTCCCAAGAGAG 6840
|||||

Dh 7121 GAAGAAAGAGCAGCTTGAGCCTTGATACATCGAATATATGTCTCCCAAGAGAGA 7180
Qy 6841 TTCCACAGCTTTACCGGCTTGAGCAGCGCTGATTTACAACCCAGCTTGTGAATG 6900
Dh 7181 TTCCACAGCTTTACCGGCTTGAGCAGCGCTGATTTACAACCCAGCTTGTGAATG 7240
Qy 6901 TGGAAAGGCGAGATTACCAACCGGCACTGTTGCGGCTGTGCTCTCTCTCTAG 6960
Dh 7241 TGGAAAGGCGAGATTACCAACCGGCACTGTTGCGGCTGTGCTCTCTCTCTAG 7300
Qy 6961 AAAACCCGACGCTCCCCAAGAGAGCGCGGACAGTGGGCTTAAGTAGAGCTCCATA 7020
Dh 7301 AAAACCCGACGCTCCCCAAGAGAGCGCGGACAGTGGGCTTAAGTAGAGCTCCATA 7360
Qy 7021 GGAGATGCTCTTCAAGAGCTGAGCCATTAAGTCTTTGCGCAGCCCCCCCAAGCGCGAT 7080
Dh 7361 GAGATGCTCTTCAAGAGCTGAGCCATTAAGTCTTTGCGCAGCCCCCCCAAGCGCGAT 7420
Qy 7081 TCGAGCTTTCCACGCGGCGCGGCGCTGCGGATTTCCGCGAGTGAAGCGCTCTGATAG 7140
Dh 7421 TCGAGCTTTCCACGCGGCGCGGCGAGCGAGTTCCGCGAGTGAAGCGCTCTGATAG 7480
Qy 7141 TTGGCCCTTTCCGAGACAGGTTTCATCTCTTTCATGCCCCCTCGAGGGGAGCTTGA 7200
Dh 7481 TTGGCCCTTTCCGAGACAGGTTTCATCTCTTTCATGCCCCCTCGAGGGGAGCTTGA 7540
Qy 7201 GATTCAGACCTGAGGCTGAGCAGTGAAGCCCCAACCCCCCGGAGGGGGGTGGA 7260
Dh 7541 GATTCAGACCTTGAAGCTGAGCAGTGAAGCTTCAACCTCCCCGAGGGGGGGGTGGA 7600
Qy 7261 GCTCCCGGCTCGGACTCGGGGCTCTGTGTTACTTTGCTCGAGAGAGACGACTCGGTG 7320
Dh 7601 ACCCCGCGCTCAGGCTCGGGGCTTGTGTACTTGTGCTCCGAGAGAGACGACTCGGTG 7660
Qy 7321 TGTGCTCCATGTCATCTCTGGAACCGGGCTCTAATACTCTGTAGTCCGGAAG 7380
Dh 7661 TGTGCTCCATGTCATCTCTGGAACCGGGCTCTAATACTCTGTAGTCCGGAAG 7720
Qy 7381 GAGAGTTACCGATTACCCCTTGAAGCACTCCCTGTGCGATATACAAAGAGTGTAC 7440
Dh 7721 GAAAGTTGCGAATTGGCCCCCTTGAAGCACTCCCTGTGCGATATACAAAGAGTGTAC 7780
Qy 7441 TGTACCAACAAGAGGCGCTCACTAAGGCTTAAAAAGTTAACTTTGATAGATGCA 7500
Dh 7781 TGTACCAACAAGAGGCGCTCACTAAGGCTTAAAAAGTTAACTTTGATAGATGCA 7840
Qy 7501 GTGCTGAGCTCTAACAAGCTGAGTCTTAAGAGCACTTAAGCTAGGCGCTCAAGTTC 7560
Dh 7841 GCGCTGAGCTCTAACAAGCTGAGTCTTAAGAGCACTTAAGCTAGGCGCTCAAGTTC 7900
Qy 7561 ACCGCAAGGCTCTCTCACTAGAGAGGCTTGCAGTTAACCCCACTTCGCAAGA 7620
Dh 7901 ACCGCAAGGCTCTCTCACTTGAAGAGGCTTGCAGTTAACTCAACCCCACTTCGCAAGA 7960
Qy 7621 TCTAATATATGGGTTTGGGGCTTAAGAGAGTCCGCACTTTCGCGAGGCGCTTAAACAC 7680
Dh 7961 TCTAATATATGGGTTTGGGGCTTAAGAGAGTCCGCACTTTCGCGAGGCGCTTAAACAC 8020
Qy 7681 ATCAAGTCCGTGGAAGAGCTTCCTGAGAGACTCAGAAAAACCAATTCCCAACCATT 7740
Dh 8021 ATCAAGTCCGTGGAAGAGCTTCCTGAGAGACTCAGAAAAACCAATTCCCAACCATT 8080
Qy 7741 ATGGCCAAAATAGAGTGTCTGTGTGAGACCCCAACAAGGGGGCAAGAAAGAGCTGCG 7800
Dh 8081 ATGGCCAAAATAGAGTGTCTGTGTGAGACCCCAACAAGGGGGTAAGAAAGAGCTGCG 8140
Qy 7801 CTTATGCTTTACCTGAGCTTCGCGGTAGAGGTCTGCGAAGATGGCCCTTTATGACATT 7860
Dh 8141 CTTATGCTTTACCTGAGCTTCGCGGTAGAGGTCTGCGAAGATGGCCCTTTATGACATT 8200
Qy 7861 ACACAAAAATTCTCTCAGGCGGTGATGGGGGCTTCTTATGATTCACATTTACCCCGCT 7920
Dh 8201 ACACAAAAATTCTCTCAGGCGGTGATGGGGGCTTCTTATGATTCACATTTACCCCGCT 8260

Qy 7921 CAGCGGTTAGATTCTCTTGAAGCATGGGCGGAAAAAGAGACCTTATGGGTTTTGCG 7980
Dh 8261 CAGCGGTTAGATTCTCTTGAAGCATGGGCGGAAAAAGAGACCTTATGGGTTTTGCG 8320
Qy 7981 TATGATACCGAGTCTTGAATCAACCGTCACTGAGAGACATCAGACTAGAGAGTCC 8040
Dh 8321 TATGATACCGAGTCTTGAATCAACCGTCACTGAGAGACATCAGACTAGAGAGTCC 8380
Qy 8041 ATATATCGGGGCTGCTCTTTCGCGAGAGGCGCACACTGGCATCACTCGCTAACTGAG 8100
Dh 8381 ATATATCGGGGCTTGTTCCTTTCGCGAGAGGCGCACACTGGCATCACTCGCTAACTGAG 8440
Qy 8101 AGACTTTACGTGAGAGGCTTATGTTCAACAGCAAGGCGCAAACTTCGAGGTACAGCGCT 8160
Dh 8441 AGACTTTACGTGAGAGGCGCTTATGTTCAACAGCAAGGCGCAAGCTTCGAGGTACAGCGCT 8500
Qy 8161 TGCCTGCGCAGCGGGGTGCTCACTCACTAGACATGGGGAACACATCACTGCTAGAA 8220
Dh 8501 TGCCTGCGCAGCGGGGTGCTTACCACTAGACATGGGGAACACATCACTGCTAGAA 8560
Qy 8221 GCCTTAGCGGCTTGTAAAGCTGACAGGATATTCGCGCCCAATGCTGTATGCGGCGAT 8280
Dh 8561 GCCTTAGCGGCTTGTAAAGCTGACAGGATATTCGCGCCCAATGCTGTATGCGGCGAT 8620
Qy 8281 GACTTGTGTATCTCAGAAAAGCAAGGGACCGAGAGAGAGAGCGAACTTGAGAGCC 8340
Dh 8621 GACTTGTGTATCTCAGAGAGCCAGGGGACCGAGAGAGAGAGCGAACTTGAGAGCC 8680
Qy 8341 TTACAGGAGGCTTATGACCGAGTATTTGCGCCCTCTGTGTGAGACCCCGGACCGAGAT 8400
Dh 8681 TTACAGGAGGCTTATGACCGAGTATTTGCGCCCTCTGTGTGAGACCCCGGAGAT 8740
Qy 8401 GATCTGAGCTGATTAACATCTTGTCTCTCAAAATGTTCTGTGAGCGCTGAGCCCAAGAGC 8460
Dh 8741 GATCTGAGCTGATTAACATCTTGTCTCTCAAAATGTTCTGTGAGCGCTGAGCCCAAGAGC 8800
Qy 8461 GCGCGAGATACTACTGACAGAGACCTTACCACTCAATGCGCGGCTGCTGGGAA 8520
Dh 8801 GCGCGAGATACTACTGACAGAGACCTTACCACTCAATGCGCGGCTGCTGGGAA 8860
Qy 8521 ACAGTTAGACATCCCTCTGCAATTCATAGGCTGGGAAACATATCAGTACGCGCGGAC 8580
Dh 8861 ACAGTTAGACATCCCTCTGCAATTCATAGGCTGGGAAACATATCAGTACGCGCGGAC 8920
Qy 8581 ATATGGGCTGCGATGGTCTGATGACACACTTCTTCCATTCTCATGCTCAAGACAG 8640
Dh 8921 ATATGGGCTGCGATGGTCTGATGACACACTTCTTCCATTCTCATGCTCAAGATACT 8980
Qy 8641 CTGGACAGAACTTCACTTTGAGATGTACGAGCGGTGACTCGGTAGTCCCTTGGAC 8700
Dh 8981 CTGGACAGAACTTCACTTTGAGATGTACGAGCGGTGACTCGGTAGTCCCTTGGAC 9040
Qy 8701 CTCCGACCTAATTTGAAGGTTTACATGGGCTTGAAGCTTTTCTCTGACATACACT 8760
Dh 9041 CTCCGACCTAATTTGAAGGTTTACATGGGCTTGAAGCTTTTCTCTGACATACACT 9100
Qy 8761 CCCCAAGACTGACAGGCTGAGCTTCAAGCCCTCAGAAAACTTGGGCGGCAACCCCTCAGA 8820
Dh 9101 CCCCAAGACTGACAGGCTGAGCTTCAAGCCCTCAGAAAACTTGGGCGGCAACCCCTCAGA 9160
Qy 8821 GCGTGAAGAGCCGAGCAGTGCAGTCAAGGCGTCCCTCATCTCCGTGGGGGAGAGCG 8880
Dh 9161 GCGTGAAGAGCCGAGCAGTGCAGTCAAGGCGTCCCTCATCTCCGTGGGGGAGAGCG 9220
Qy 8881 GCGTGTGGGTGAGTATCTCTTCAATGGGGGCTGGAAGAACCAAGTCAAACTCACTCA 8940
Dh 9221 GCGTGTGGGTGAGTATCTCTTCAATGGGGGCTGGAAGAACCAAGTCAAACTCACTCA 9280
Qy 8941 TTGCGGAGAGCGGCTCTCTGATTTATCAGCTGGTTACCGTCCGCGCGGCGGAGC 9000
Dh 9281 TTGCGGAGAGCGGCTCTCTGATTTATCAGCTGGTTACCGTCCGCGCGGCGGAGC 9340

Qy 9001 GACATTATATACAGACGCTGTCGCGCCGACCCCGCTTATATGCTTTGGCCCTACTCTTA 9060
|
|
|
Db 9341 GACATTATATACAGACGCTGTCGCGCCGACCCCGCTTATATGCTTTGGCCCTACTCTTA 9400
|
|
|
Qy 9061 CTTTGTGTAGGGGTAGGCTTTTCTTACTCTCCCGCTCGG 9099
|
|
|
Db 9401 CTTTGTGTAGGGGTAGGCTTTTCTTACTCTCCCGCTCGG 9439
|
|
|
RESULT 7
E07362 9589 bp RNA linear PAT 29-SEP-1997
LOCUS
DEFINITION CNA of Hepatitis non-A non-B virus.
ACCESSION E07362
VERSION E07362.1 GI:2175501
KEYWORDS JP 1994121689-A/2.
SOURCE
ORGANISM
unidentified
unclassified.
1 (bases 1 to 9589)
REFERENCE
AUTHORS Okamoto H. and Nakamura T.
TITLE NON-A NON-B HEPATITIC VIRUS GENE, POLYNUCLEOTIDE, POLYPEPTIDE,
ANTIGEN AND ANTIBODY DETECTION SYSTEM
JOURNAL Patent: JP 1994121689-A 2 06-MAY-1994;
NAKAMURA TETSUO
COMMENT OS Hepatitis non-A non-B virus
PN JP 1994121689-A/2
PD 06-MAY-1994
PF 09-AUG-1991 JP 1991287402
PI OKAMOTO HIROAKI, NAKAMURA TETSUO
PC C12N15/51, A61K39/29, A61K39/395, A61K39/395, C07K13/00, C12P21/02,
PC C12P21/08,
PC G01N33/53, G01N33/576, G01N33/577//A61B10/00;
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1. 9589
FT /strain='HCV-J6'.
FEATURES
source 1. 9589
Location/Qualifiers
/organism='unidentified'
/mol_type='genomic RNA'
/db_xref='taxon:32644'
ORIGIN
Query Match 93.4%: Score 8494.2; DB 6; Length 9589;
Best Local Similarity 95.8%: Pred. No. 0;
Matches 8721; Conservative 0; Mismatches 378; Indels 0; Gaps 0;
Qy 1 ATGAGCAAAATCTTAACCTCAAGAAAGAAACCAAAAGAAACCAACCGTCGCCCAAA 60
|
|
|
Db 341 ATGAGCAAAATCTTAACCTCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 400
|
|
|
Qy 61 GACGTTAAGTTTCCGGGCGCGCGCGAGATGTTGGCGAGTAACTTGTGGCGCGCAGG 120
|
|
|
Db 401 GACGTTAAGTTTCCGGGCGCGCGCGAGATGTTGGCGAGTAACTTGTGGCGCGCAGG 460
|
|
|
Qy 121 GGGCCCAAGTTGGGTGGTGGCGGACAAAGAAAGCTTCGAGCGGTCCGAGCCACGTGGA 180
|
|
|
Db 461 GGGCCCAAGTTGGGTGGTGGCGGACAAAGAAAGCTTCGAGCGGTCCGAGCCACGTGGA 520
|
|
|
Qy 181 AGGCGCCAGCCCAATCCCTAAAGATCGCGCTCACTGSCAAATCTCGGGGAAACACAGGA 240
|
|
|
Db 521 AGGCGCCAGCCCAATCCCTAAAGATCGCGCTCACTGSCAAATCTCGGGGAAACACAGGA 580
|
|
|
Qy 241 TACCCCTGGCCCTTATACGGGAATGAGGAGCTGGGCTGGGAGAGATGGCTTCCTGCCCC 300
|
|
|
Db 581 TACCCCTGGCCCTTATACGGGAATGAGGAGCTGGGCTGGGAGAGATGGCTTCCTGCCCC 640
|
|
|

Qy 301 CGAAGTTCCGCTCCCTTTGGGGCCCAATGATACCCCGGAGTAAAGTGGCGCAACGTGGGT 360
|
|
|
Db 641 CGAAGTTCCGCTCCCTTTGGGGCCCAATGATACCCCGGAGTAAAGTGGCGCAACGTGGGT 700
|
|
|
Qy 361 AAGGTATCGATACCTTAACGTGCGGCTTGGCCGACTCATGGGGTACATCCCTGTGCTG 420
|
|
|
Db 701 AAGGTATCGATACCTTAACGTGCGGCTTGGCCGACTCATGGGGTACATCCCTGTGCTG 760
|
|
|
Qy 421 GGGCCCCCGCTCGGCGCGCTCGGACAGCTCTGCGCATGGCGATGGCGTGAAGTCTGGAGAG 480
|
|
|
Db 761 GGGCCCCCGCTCGGCGCGCTCGGACAGCTCTGCGCATGGCGATGGCGTGAAGTCTGGAGAG 820
|
|
|
Qy 481 GGGGTTAATTTTGAACAGGAACTTACCGGTTGCTCTTTCTATCTTTCTTCTGCGGC 540
|
|
|
Db 821 GGGGTTAATTTTGAACAGGAACTTACCGGTTGCTCTTTCTATCTTTCTTCTGCGGC 880
|
|
|
Qy 541 CTGCTGCTCTGATACCAACCCCGGTCTCGCTGCGAAGTGAAGAAATCATGATCCGGC 600
|
|
|
Db 881 CTGCTGCTCTGATACCAACCCCGGTCTCGCTGCGAAGTGAAGAAATCATGATCCGGC 940
|
|
|
Qy 601 TACATGATGATTAACGACGTCGACCAATGATAGAGATTAACCTGGAGCTCCAGGCTGCTC 660
|
|
|
Db 941 TACATGATGATTAACGACGTCGACCAATGATAGATTAACCTGGAGCTCCAGGCTGCTC 1000
|
|
|
Qy 661 CTCACAGTCCCGGAGTGCCTCCGTCGAGAAAGTGGGGAATGCATCTCAGTGGTGGATA 720
|
|
|
Db 1001 CTCACAGTCCCGGAGTGCCTCCGTCGAGAAAGTGGGGAATGCATCTCAGTGGTGGATA 1060
|
|
|
Qy 721 CCGGCTCACCGAATGTCGCGAGCCGTCGAGCGGCGCCGCTCAGCGAGGCTTGGAGCG 780
|
|
|
Db 1061 CCGGCTCACCGAATGTCGCGAGCCGTCGAGCGGCGCCGCTCAGCGAGGCTTGGAGCG 1120
|
|
|
Qy 781 CACATGACATGTTGATGATGTCGCGACGCTCTGCTCTGCGCTTACAGTGGGGAGCTC 840
|
|
|
Db 1121 CACATGACATGTTGATGATGTCGCGACGCTCTGCTCTGCGCTTACAGTGGGGAGCTC 1180
|
|
|
Qy 841 TGGCGTGGGGTGAATGCTCGAGCCCAATGTTGATGTCGCGCGAGCACTAGTGT 900
|
|
|
Db 1181 TGGCGTGGGGTGAATGCTCGAGCCCAATGTTGATGTCGCGCGAGCACTAGTGT 1240
|
|
|
Qy 901 GTCCAGAGCTGCAATGCTCTCATCTACCTGTGTAACATGACAGCAACCGCATGGAGTGG 960
|
|
|
Db 1241 GTCCAGAGCTGCAATGCTCTCATCTACCTGTGTAACATGACAGCAACCGCATGGAGTGG 1300
|
|
|
Qy 961 GACATGATGATGAATGATGTCGCGACGCTCAACCATGATCTTGGGCTGACGAGTGGCTG 1020
|
|
|
Db 1301 GACATGATGATGAATGATGTCGCGACGCTCAACCATGATCTTGGGCTGACGAGTGGCTG 1360
|
|
|
Qy 1021 CCGGAGGTCATTAAGCATCATTAAGCGGGGCTCATTTGGGGCGTCATGTCGGCTTGACC 1080
|
|
|
Db 1361 CCGGAGGTCATTAAGCATCATTAAGCGGGGCTCATTTGGGGCGTCATGTCGGCTTGACC 1420
|
|
|
Qy 1081 TACTTCTTAAGAGGAGGCGTGGGCGAAAGTGTGTCTATCTTCTGTTGGCGCGGG 1140
|
|
|
Db 1421 TACTTCTTAAGAGGAGGCGTGGGCGAAAGTGTGTCTATCTTCTGTTGGCGCGGG 1480
|
|
|
Qy 1141 GTGAGCGCGGCAACCATATCTGTTGGGGGTTCTGCGCGGAGACACCGGGCGCTTACC 1200
|
|
|
Db 1481 GTGAGCGCGGCAACCATATCTGTTGGGGGTTCTGCGCGGAGATTAACGACAGACCTTACC 1540
|
|
|
Qy 1201 AGCTTAATTTGACATGGGCGCCGAGGAGAAATCCAGCTGCTTAACCAATGGGAGCTGG 1260
|
|
|
Db 1541 AGCTTAATTTGACATGGGCGCCGAGGAGAAATCCAGCTGCTTAACCAATGGGAGCTGG 1600
|
|
|
Qy 1261 CACATCAACCGGACCGGCTGAACTGCAATGACTCTTGGACACCGGCTTTATGCGTCT 1320
|
|
|
Db 1601 CACATCAACCGGACCGGCTGAACTGCAATGACTCTTGGACACCGGCTTTATGCGTCT 1660
|
|
|
Qy 1321 CTGTTCTACACCGGACGCTTCACTGTCAGAGATGTCGGAAGGACATGTCGGCTGCGCG 1380
|
|
|
Db 1661 CTGTTCTACACCGGACGCTTCACTGTCAGAGATGTCGGAAGGACATGTCGGCTGCGCG 1720
|
|
|
Qy 1381 AGTATGAGGCTTTCGGGTGGAGTGGGGCGCTTGCATATGAGGATTAATGTCACCAAT 1440
|
|
|

Db 1721 AGATATGAGGCTTTGGGGGATGCGGCGCTTCAATATGAGACATGTCACCAAT 1780
Qy 1441 CCAGAGATATGAGACCCATATGCTGGCACTTACCCACCAAGGAGTGGCGGTCTCC 1500
Db 1781 CCAGAGATATGAGACCGTATTTGCTGGCACTTACCCACCAAGAGAGTGGTATGTCCTC 1840
Qy 1501 GCGAGACCTGTGTGGGCCAGTGTACTGTTCACCCCGAGCCAGTGTATGGGACG 1560
Db 1841 GCGAGCTGTGTGGGCCAGTGTACTGTTCACCCCGAGCCAGTGTATGGGATG 1900
Qy 1561 ACCGACAGGCTTGGAGGCGCCACTTACAGTGGGGGAGATGAGACAGTGTCTTCTA 1620
Db 1901 ACCGATAGACTTGGAGGCGCCACTTACAGTGGGGGAGATGAGATGATCTTCTA 1960
Qy 1621 TTGAACGACCTGACACACCGCTGGGGTCAATGGTGGCTGACGATGAGTAACTTCTT 1680
Db 1961 TTGAACGACCTGACACACCGCGAGGGGTCAATGGTGGCTGACGATGAGTAACTTCACT 2020
Qy 1681 GGGTACACCAAGACTTGGCGGACCAACCTGCGGTACTAGAGCTGACTTCAAGCGCAGC 1740
Db 2021 GGGTACACCAAGACTTGGCGGACCAACCTGCGGTACTAGAGCTGACTTCAATGCGAGC 2080
Qy 1741 ACCGACTGTGTGGGCCAGGACTGTTTATGAGAGATCTGTATACCACTTACCTCAAA 1800
Db 2081 ATGACTTGTGTGGGCCAGGACTGTTTATGAGAGATCTGTATACCACTTACCTCAAA 2140
Qy 1801 TGGGCTCTGGGCGCTGGCTCAAGCCCAAGGTCGTATGACTACCCCTACAGGCTCTGG 1860
Db 2141 TGGGCTCTGGGCGCTGGCTCAAGCCCAAGGTCGTATGACTACCCCTACAGGCTCTGG 2200
Qy 1861 CATTAACCTGCAACAGTTAACTATACCATCTTCAAAATGAATGATGTGGAGGGGT 1920
Db 2201 CATTAACCTGCAACAGTTAACTATACCATCTTCAAAATGAATGATGTGGAGGGGT 2260
Qy 1921 GAGCAGAGGCTCAAGGTCATGCAATTTTCACTGTGGGATGCTTGGCACTTGGAGAGC 1980
Db 2261 GAGCAGAGGCTCAAGGTCATGCAATTTTCACTGTGGGATGCTTGGCACTTGGAGAGC 2320
Qy 1981 AGAGACAGAGTCAACGTCTCTTGTGTGCACTCCACAGGATGGGCACTTTTAACT 2040
Db 2321 AGAGACAGAGTCAACGTCTCTTGTGTGCACTCCACAGGATGGGCACTTTTAACT 2380
Qy 2041 TGTCTTACTCGAGCTGCCCGCTGTGTGCACTGTGTCTTCCACTTCCACAAACATC 2100
Db 2381 TGTCTTACTCGAGCTGCCCGCTGTGTGCACTGTGTCTTCCACTTCCACAAACATC 2440
Qy 2101 GTGAGCTATCAATTCATGTATGGCTTATCACCTGCCCTCAAAATACATGTCGATGG 2160
Db 2441 GTGAGCTATCAATTCATGTATGGCTTATCACCTGCCCTCAAAATACATGTCGATGG 2500
Qy 2161 GAGTGGGTAAATCTTATTCCTGCTCTTATGCGGACGCGAGGTTTGGCGCTGTATGG 2220
Db 2501 GAGTGGGTAAATCTTATTCCTGCTCTTATGCGGACGCGAGGTTTGGCGCTGTATGG 2560
Qy 2221 ATGTCTATCTTGTGGGCGAGCGGAGAGCACTAGAGAAAGCTGATCATCTTGGACGT 2280
Db 2561 ATGTCTATCTTGTGGGCGAGCGGAGAGCACTAGAGAAAGCTGATCATCTTGGACGT 2620
Qy 2281 GCGAGCGCAGTACGTGCAATGGCTTCTATATTTTGTCACTTTTTCGTGGCTGTGG 2340
Db 2621 GCGAGCGCAGTACGTGCAATGGCTTCTATATCTTGTCACTTTTTCGTGGCTGTGG 2680
Qy 2341 TACATCAAGGGTGGGGATGTCCTTATGACTATTCCTCACTGGCTGTGGTCTTT 2400
Db 2681 TACATCAAGGGTGGGGATGTCCTTATGACTATTCCTCACTGGCTGTGGTCTTT 2740
Qy 2401 AGCTTACTGTCTTAGCATTTGCCCAACAGGCTTATGCTTATGACGATCTGTGATGGC 2460
Db 2741 GGCCTACTGTCTTAGCATTTGCCCAACAGGCTTATGACGATCTGTGATGGT 2800
Qy 2461 CAGATAGAGCGGCTGTGCTGTATATCACTCTTCTTACTCTACCCCCGGTATTAAG 2520

Db 2801 CAGATAGAGCACTCTGTGTACTGATCACTCTCTTTACACTACCCCCGGGTATTAAG 2860
Qy 2521 ACCCTTCTGAGCGGGTTTTGTGTGTGTGTACTATCTTGTGACCTTGGGGAGGCTATG 2580
Db 2861 ACCCTTCTGAGCGGGTTTTGTGTGTGTGTGTACTATCTTGTGACCTTGGGGAGGCTATG 2920
Qy 2581 GTCCAGAGTGGGCAACCACTTATGCAAGTGGCGGGTGGCGGTATGGCATCATATGGGC 2640
Db 2921 GTCCAGAGTGGGCAACCACTTATGCAAGTGGCGGGTGGCGGTATGGCATCATATGGGC 2980
Qy 2641 GTCCCATATTCATCCAGGTTGGGTGTGATATACCAAGGCTTGGCGGGTCTT 2700
Db 2981 GTCCCATATTCATCCAGGTTGGGTGTGATATACCAAGGCTTGGCGGGTCTT 3040
Qy 2701 GGGCTGTCTTACCTCTTAAAGGTGCTTGAAGCGCGTCCGATCTTGTCAAGGCTCAC 2760
Db 3041 GGGCTGTCTTACCTCTTAAAGGTGCTTGAAGCGCGTCCGATCTTGTCAAGGCTCAC 3100
Qy 2761 GCTTACTGAGATGTGACCAATGGCAGAGCATCTGCGGGGGGAGATAGTCCAGATG 2820
Db 3101 GCTTACTGAGATGTGACCAATGGTAAAGCATCTGCGGGGGGTAGTACGTCCAGATG 3160
Qy 2821 GGGCTACTAGCCTTGGGAGGGTGGACTGGCACTTATGATGACCACTACCCCTATG 2880
Db 3161 GTGCTACTAGCCTTGGGAGGGTGGACTGGCACTTATGATGACCACTACCCCTATG 3220
Qy 2881 TGGGATTTGGGCTGTAGTGGCTGCGGAGCTTGGCGGTGCGGTTAGGCTTATCTTC 2940
Db 3221 TGGGATTTGGGCTGTAGTGGCTGCGGAGCTTGGCGGTGCGGTTAGGCTTATCTTC 3280
Qy 2941 AGTCCATGAGAGAAAGTCAATTTGTGTGGGAGCGGAGACAGCTCTTGTGGGACAT 3000
Db 3281 AGTCCATGAGAGAAAGTCAATTTGTGTGGGAGCGGAGACAGCTCTTGTGGGAGATTC 3340
Qy 3001 TTACACGACTTCCCGGTGTCGCGCCGACTTGGTGGGAGGCTCTTGGAGCCAGCTGAT 3060
Db 3341 TTACACGACTTCCCGGTGTCGCGCCGACTTGGTGGGAGGCTCTTGGAGCCAGCTGAT 3400
Qy 3061 GGTATATCTCCAGAGGGGTGAGTCTTCTGCGCCCATCATGCTTATCGCCAGCAGACA 3120
Db 3401 GGTATATCTCCAGAGGGGTGAGTCTTCTGCGCCCATCATGCTTATCGCCAGCAGACA 3460
Qy 3121 CGTGCCCTTTTGGGACCAATAGTGTGAGATGACGCGGCGGACCAAGACAAACAGGCT 3180
Db 3461 CGTGCCCTTTTGGGACCAATAGTGTGAGATGACGCGGCGGACCAAGACAAACAGGCT 3520
Qy 3181 GGGGAAATCAAGTCTGTGTCACAGTCACTGAGTCTTCTGCGGAAATCATCATCTGGGG 3240
Db 3521 GGGGAAATCAAGTCTGTGTCACAGTCACTGAGTCTTCTGCGGAAATCATCATCTGGGG 3580
Qy 3241 GTTTGTGACTGTCTTACCATGAGCTGGCAACAGACTTGGCCGGCTTCAAGGGTCCG 3300
Db 3581 GTTTGTGACTGTCTTACCATGAGCTGGCAACAGACTTGGCCGGCTTCAAGGGTCCG 3640
Qy 3301 GTTACGCAATGTATCTCCAGTGTGAGGGGACTTATGATGGTGGCCAGCCCTTGGG 3360
Db 3641 GTTACGCAATGTATCTCCAGTGTGAGGGGACTTATGATGGTGGCCAGCCCTTGGG 3700
Qy 3361 ACTAAATCTTTGAGGCGGTCAGATGATGAGGCGGTGACCTGTACTGTGACGCGGAA 3420
Db 3701 ACCAAATCTTTGAGGCGGTCAGATGATGAGGCGGTGACCTGTACTGTGACGCGGAA 3760
Qy 3421 GCTGATGTATCCCGGCTGGAAGACCGCGGAGCAAAACCGGAGAGGCTACTCTCCGAGA 3480
Db 3761 GCTGATGTATCCCGGCTGGAAGACCGCGGAGCAAAACCGGAGAGGCTACTCTCCGAGA 3820
Qy 3481 CTTCTTTCACTTTGAAGGGGTCTTCAAGAGGCGGCTGTATGCCCCAGGGGCAAGCT 3540
Db 3821 CTTCTTTCACTTTGAAGGGGTCTTCAAGAGGCGGCTGTATGCCCCAGGGGCAAGCT 3880
Qy 3541 GTGAGTGTCTTCCGGGAGCTGTGTGCTCTCCGGGAGGCTGAGCTTAAGTCATATGATTC 3600
Db 3881 GTGAGTGTCTTCCGGGAGCTGTGTGCTCTCCGGGAGGCTGAGCTTAAGTCATATGATTC 3940

3601 CCGGTTGAGACACTCGACATGCTCAACGCGGTCGCCCACTTTAGTGAACAAGACACCA 3660
3941 CCGGTGAGACACTGTGACATGCTCACTCGGTCCGCCACTTTAGTGAACAAGACACCA 4000
3661 CTTGTGTGCCCCAGACTATCAAGTGTGGGTACTTGGCATGCCCGACTGGCAGTGAAG 3720
4001 CTTGTGTGCCCCAACTTATCAAGTGTGGGTACTTGGCATGCCCGACTGGTATGGAAG 4060
3721 AGACCAAAAGTTCTGTGCGATATGCTGTCAGGGGTATAAAGTCTAGTCTTAATCC 3780
4061 AGACCAAAAGTCTGTGCGATATGCTGTCAGGGGTATCAAAAGTCTAGTCTTAATCC 4120
3781 TCAGTGGCTGCAACCTGGGGGTTTGGGGCGTACTGTGCTAAAGGACAATGGCATCAATCC 3840
4121 TCGGTGGCTGCAACCTGGGGGTTTGGGGCGTACTGTGCTAAAGGACAATGGCATCAATCC 4180
3841 AACATTAGGACTGAGTCAAGACTGTGACGACCGGGGCGCCATCACTGCTCCACATAT 3900
4181 AACATTAGGACTGAGGGGTCAAGACTGTGACGACCGGGGCGCCATCACTGCTCCACATAT 4240
3901 GGCAAATTCCTCGCCGATGGGGGCTGTGCGGGCGGCGCTTACGACATCATATATGAT 3960
4241 GGCAAATTCCTCGCCGATGGGGGCTGTGCGGGCGGCGCTTATGACATCATATATGCGAT 4300
3961 GAATGCGATGCGGTGACCTACACCACTCTTGGCATGGAAACAGTCTTGAATCAAGCA 4020
4301 GAATGCGATGCGGTGACCTACACCACTCTTGGCATGGAAACAGTCTTGAATCAAGCA 4360
4021 GAGACAGCTGGGGTCAAGACTTACTGCTGCTTACAGCTACAGCTACGCCCCCTGGGTCAGTACA 4080
4361 GAGACAGCTGGGGTCAAGGGCTTAACTGTAAGCTTACAGCTACGCCCCCTGGGTCAGTACA 4420
4081 ACCCCCCCAACCAATAGAGAGTGGCCCTTGGGAGAGAGAGGCGAGATCCCTTCTAT 4140
4421 ACCCCCCCAACCAATAGAGAGTGGCCCTTGGGAGAGAGAGGCGAGATCCCTTCTAT 4480
4141 GGGAGGGCGATTCCCGTGTCTTACATCAAGAGAGAGAGATCTTCTGCAATTA 4200
4481 GGGAGGGCGATTCCTGTATACATCAAGAGAGAGAGATCTTCTGCAATTA 4540
4201 AAGAAAAAGTGTACAGAGCTCGCGCGGCGCTTGGGGGTATGGGCTTGAATCAAGTGGCA 4260
4541 AAGAAAAAGTGTACAGAGCTCGCGCGGCGCTTGGGGGTATGGGCTTGAATCAAGTGGCA 4600
4261 TACTACAGAGGTTGAGCTGCTCCGTAATCAACTCAAGAGAGAGCTGATGCTGCGCC 4320
4601 TACTACAGAGGTTGAGCTGCTCCGTAATCAACTCAAGAGAGAGCTGATGCTGCGCC 4660
4321 ACCGAGCCCTCAATGACAGGGTATACCTGGGAGCTTGAATCGGTGATCGATCAAGCTA 4380
4661 ACCGAGCCCTCAATGACAGGGGTTTACTGGAGACTTGGATCGGTGATCGATCAAGCTA 4720
4381 GCGGTCACTCAAGTTGTAGACTTCAAGTTAGACCCCACTTCAACATCAAGAGAT 4440
4721 GCGGTCACTCAAGTTGTAGACTTCAAGTTAGACCCCACTTCAACATCAAGAGAT 4780
4441 GTCCCTCAAGAGCTGTCTCACTGACGACGCGGGGCTGCAACGAGTGAAGGAGCTG 4500
4781 GTCCCTCAAGAGCTGTCTCACTGACGACGCGGGGCTGCAACGAGTGAAGGAGCTG 4840
4501 GGCATTATATAGTATGTTTCACTGAGTGAAGGCTCAAGAAATGTTTGAACAGTATG 4560
4841 GGCATTATATAGTATGTTTCACTGAGTGAAGGCTCAAGAAATGTTTGAACAGTATG 4900
4561 CTCTGTGAGTGTACGACGAGGGGCGCAATGATAGCTCAACCATCGAGAGACCA 4620
4901 CTCTGTGAGTGTACGAGTGAAGGGGCGCAATGATAGCTCAACAGAGAGACCA 4960
4621 GTCAAGCTCAAGGCGTATTTCAACAGCGCGGTTTGTCTGTGTGCAAGACATCTT 4680
4961 GTCAAGCTCAAGGCGTATTTCAACAGCGGTTTGTCTGTGTGCAAGACATCTT 5020

4681 TTTTGGAGGCAAGTTTTCACCGGCTCAACACATAGATGCCACTTCCCTTCCAAACA 4740
5021 TTTTGGAGGCAAGTTTTCACCGGCTCAACACATAGATGCCACTTCCCTTCCAAACA 5080
4741 AAGCAATCGGGGAAAAATTTTCGATATTAACAGCTTACAGGCTACAGTGTGCGT 4800
5081 AAGCAATCGGGGAAAAATTTTCGATATTAACAGCTTACAGGCTACAGTGTGCGT 5140
4801 GCCAAAGCCCCCCCCCTGCTGGGACGTCATGTGGAAAGTTTGAATCTGCACTCAAGCC 4860
5141 GCCAAAGCCCCCCCCCTGCTGGGACGTCATGTGGAAAGTTTGAATCTGCACTCAAGCC 5200
4861 ACACTGTGGGCCCCACACTCTCTGTAACCGCTTGGGCTGTGTTACCAAGAGTCA 4920
5201 ACACTGTGGGCCCCACACTCTCTGTAACCGCTTGGGCTGTGTTACCAAGAGTCA 5260
4921 CTCACATCTCCGTGACGAATATCATCGCCACTGTGACAGCCGACTTGAAGTATG 4980
5261 CTCACATCTCTGTGACGAATATCATCGCCACTGTGACAGCCGACTTGAAGTATG 5320
4981 ACCAGCATAGGTTCTTGGCAGGGGAGTCTTGGCGGCGCTCGCCGCTATTCCTG 5040
5321 ACCAGCATAGGTTCTTGGCAGGGGAGTCTTGGCGGCGCTCGCCGCTATTCCTG 5380
5041 ACCGGGTGTTTGCATCATCGGCGCTTGCACATTAACAGAGAGCCGCTGTGCGCG 5100
5381 ACCGGGTGTTTGCATCATCGGCGCTTGCACATTAACAGAGAGCCGCTGTGCGCG 5440
5101 GACAAAGAGTCTCTATAGAGGCTTTTGAATGAGATGAGAAATGTGCTCTTGA 5160
5441 GACAAAGAGTCTCTATAGAGGCTTTTGAATGAGATGAGAAATGTGCTCTTGA 5500
5161 CTGATTGAAGAGGGGAGGAGGAGATGCGGAGATGCTGAAGTCCAAATCCAAAGCT 5220
5501 CTGATTGAAGAGGGGAGGAGGAGATGCGGAGATGCTGAAGTCCAAAGCT 5560
5221 CAGCAAGCTTCCAAAGGCTCAAGACATACAAACCACTGTGAGGCTTCAATGAGCC 5280
5561 CAGCAAGCTTCCAAAGGCTCAAGACATACAAACCACTGTGAGGCTTCAATGAGCC 5620
5281 GTAGAACAATCTGAGGCGAAACACATGTGGAATCTTATAGCGGATCCAAATCT 5340
5621 GTAGAACAATCTGAGGCGAAACACATGTGGAATCTTATAGCGGATCCAAATCT 5680
5341 GGAATATCAACATCTGCGAGGAAACCTGCAAGTATGCTTCAATGAGGCTCAAGT 5400
5681 GGAATATCAACATCTGCGAGGAAACCTGCAAGTATGCTTCAATGAGGCTCAAGT 5740
5401 CTCACAGTCCGCTGTCACAAAGCAACATATCTTCTCAACATTTTGGGGGCTGGCT 5460
5741 CTCACAGTCCGCTGTCACAAAGCAACATATCTTCTCAACATTTTGGGGGCTGGCT 5800
5461 GCATCCCAATTTGACACACCCGCGGGGCACTGGGCTTGTGTCAGTGGCTTATG 5520
5801 GCATCCCAATTTGACACACCCGCGGGGCACTGGGCTTGTGTCAGTGGCTTATG 5860
5521 GCTGCGGTAGAGATATAGCTTATAGTAAAGTCTATGAGACATCTCTGAGAGGAT 5580
5861 GCTGCGGTAGAGATATAGCTTATAGTAAAGTCTATGAGACATCTCTGAGAGGAT 5920
5581 GCGGGAATTTGCGGGGCTCGTGTGATTAAGATCATGCTGGGAGAGAGCCCTCAT 5640
5921 GCGGGAATTTGCGGGGCTCGTGTGATTAAGATCATGCTGGGAGAGAGCCCTCAT 5980
5641 GAGGATGTGCTCAACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5700
5981 GAGGATGTGCTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6040
5701 ATCTGCGCGGCAATTTGCGGCGGACAGTGGGAGCGGGGAGAGCCGCTCAATGAG 5760
6041 ATCTGCGCGGCAATTTGCGGCGGACAGTGGGAGCGGGGAGAGCCGCTCAATGAG 6100
5761 AATGACATCAATTCCTTGTCTCAAGAGAAATCAAGTGGCCCCCACTAGTACG 5820

Db 6101 AATAGACTCATGCTCTTGTGCTTCAGAGGAAACACAGTGCCTCCCAACCCACTAGAGG 6160
Qy 5821 GATCGGATGCGCTGCGAGCGCTGTGACCCAACTACTTGGCTCCCTTAACCAATTAACACCTG 5880
Db 6161 GATCGGATGCGCTGCGAGCGCTGTGACCCAACTACTTGGCTCCCTTAACCAATTAACACCTG 6220
Qy 5881 CTGAGAACTCCAGCACTGATTAATGAGAGCTGCCCCATCCCATGCGCGCTGTGG 5940
Db 6221 CTGAGAGACTCCAGCACTGATTAATGAGAGCTGCCCCATCCCATGCGCGCTGTGG 6280
Qy 5941 CTGCGGATGTGTGGAGCTGGGTTTGACCATCTTAACAGACTTTAAAACTGCTGACC 6000
Db 6281 CTGCGGATGTGTGGAGCTGGGTTTGACCATCTTAACAGACTTTAAAACTGCTGACC 6340
Qy 6001 TCCAAATTATTCGCAAGATGCGCGCTCCCTTTGTCTCTGTAAGAGGGGTCAAG 6060
Db 6341 TCCAAATTGTTCGCAAGATGCGCTGCTCCCTTTATCTCTGTCAAGAGGGGTCAAG 6400
Qy 6061 GCGGTGTGGCCGCGCACTGGCATCATGACACACGCTGTCTTGCGCGCCAAATCTCT 6120
Db 6401 GCGGTGTGGCTGGCACTGGTATCATGACACACGCTGTCTTGCGCGCCAAATCTCT 6460
Qy 6121 GCGAATGTCCGCTTGGGCTTCATGAGATACAGGGGCTTAAGACTGECATGAATCTGTG 6180
Db 6461 GCGAATGTCCGCTTGGGCTTCATGAGAATTAAGGGGCCAAAACCTGCAATATCTGG 6520
Qy 6181 CAGGGGACCTTTCCTATGCAATGTGTAACAGGAGGCGACAGTGCCTGCGGAAACCGCGGCA 6240
Db 6521 CAGGGGACCTTTCCTATGCAATGTGTAACAGGAGGCGACAGTGCCTGCGGAAACCGCGCA 6580
Qy 6241 AACTTTAAGGTGCGCATCTGAGAGGTGCGCGCTCAGAGTACGCGAGGTGAGCGACAC 6300
Db 6581 AACTTTAAGTTCGCAATCTGAGAGGTGCGCGCTCAGAGTACGCGAGGTGAGCGACAC 6640
Qy 6301 GGGTCATACCACTACATACAGACTCACACCTGATTACTTGAAGTCCCTTGCCTA 6360
Db 6641 GGGTCATACCACTACATACAGACTCACTGATTACTTGAAGTCCCTTGCCTA 6700
Qy 6361 CCCCTCCGAGTCTTTCTGAGGTGAGAGGAGTGCATAGGTTTGCCTCCACA 6420
Db 6701 CCTTCTCAGAGTCTTTCTGAGGTGAGAGGAGTGCATAGGTTTGCCTCCACA 6760
Qy 6421 CCGAAGCGTCTTTCGCGAGTGAAGTCTCGTTCGCTTGGGCTTAATTCATTTGTCTG 6480
Db 6761 CCGAAGCGTCTTTCGCGAGTGAAGTCTCGTTCGCTTGGGCTTAATTCATTTGTCTG 6820
Qy 6481 GGGTCCAGCTTCCTTGCAGCCTGAAACCGACACAGAGTATTTGATTCATGCTAACA 6540
Db 6821 GGGTCCAGCTTCCTTGCAGTCCGAACTGACACAGAGTATTTGATTCATGCTAACA 6880
Qy 6541 GATCCATCTCATATCAAGCGGAGACTGACAGCGGGGTTTAAAGCGGGGTCACCCCA 6600
Db 6881 GATCCATCTCATATCAAGCGGAGACTGACAGCGGGGTTTAAAGCGGGGTCACCCCA 6940
Qy 6601 TCCGAGGCAAGCTCTTCGCGAGGCACTATCGGCAACCATGCTTGCAGGCACTGCAACC 6660
Db 6941 TCCGAGGCAAGCTCTTCAGCGAGGCACTATCGGCAACCATGCTTGCAGGCACTGCAACC 7000
Qy 6661 ACCGACGGCAAGCTTATGATGTGACATAGTGTGAATCTTAACCTGTTCAATGGGGGCGAT 6720
Db 7001 ACCGACGGCAAGGCTTATGATGTGACATAGTGTGAATCTTAACCTGTTCAATGGGGGCGAT 7060
Qy 6721 GTGATCGGATAGAGTCTGGGTCCAAAGGTGTGTTGTGAATCTCTCGAACCCAAATGATC 6780
Db 7061 GTGATCGGATAGAGTCTGGGTCCAAAGGTGTGTTGTGAATCTCTCGAACCCAAATGATC 7120
Qy 6781 GAAGAAAGAGAGCACTTGAAGCTTCGATACATCAGAAATACATGCTCCCAAGAGAGG 6840
Db 7121 GAAGAAAGAGAGCACTTGAAGCTTCGATACATCAGAAATACATGCTCCCAAGAGAGAG 7180
Qy 6841 TTCCCAACAGCTTACCGGCTGGGCAAGCGCTGATTAACAACCGCTTGTGAAATCG 6900

Db 7181 TTCCCAACAGCTTACCGGCTTGGGCAAGCGCTGATTAACAACCAACCGCTTGTGAAATCG 7240
Qy 6901 TGAAAAAGGCAAGATTACCAACCGGCACTGTGTGGGGGTGTGCTCTCCCTCTCTAGG 6960
Db 7241 TGAAAAAGGCAAGATTACCAACCGGCACTGTGTGGGGGTGTGCTCTCCCTCTCTAGG 7300
Qy 6961 AAAACCCGAGCGCTCCCAAGAGAGCGCGGACAGTGGGCTTAAGTAGAGACTTCATTA 7020
Db 7301 AAAACCCGAGCGCTCCCAAGAGAGCGCGGACAGTGGGCTGTGAGTAGAGCTTCATTA 7360
Qy 7021 GGAAGTGCCTTCAACAGCTGACATTAATGCTCTTGTGCAAGCCCCCAAGCGCGAT 7080
Db 7361 GGAAGTGCCTTCAACAGCTGACATTAATGCTCTTGTGCAAGCCCCCAAGCGCGAT 7420
Qy 7081 TCAGGCTTTCACAGGGGCGGGCGGTGCGGATTCGCGAGTCCAGAGCTCTGATGAG 7140
Db 7421 TCAGGCTTTCACAGGGGCGGGCGGATTCGCGAGTTCGCGAGTTCGCGAGTTCGAGTGAG 7480
Qy 7141 TTGGCCCTTTCGAGAGCAGGTTCCATCTTTCATATGCCCCCTCGAGGGGAGCTTGA 7200
Db 7481 TTGGCCCTTTCGAGAGCAGGTTCCATCTTTCATATGCCCCCTCGAGGGGAGCTTGA 7540
Qy 7201 GATCCAGACTGAGAGCTGAGAGGAGAGCCCAACCCCGCCCAAGGGGGGAGTGA 7260
Db 7541 GATCCAGACTGAGAGCTGAGAGGAGAGCTTCAACCTCCCGCCCAAGGGGGGAGTGA 7600
Qy 7261 GCTCCGAGCTCGGAGCTCGGGGTCTGTGCTACTTGTCTCGAGAGAGAGCACTCCGTG 7320
Db 7601 ACCCGGAGCTCAGGCTCGGGGTCTGTGCTACTTGTCTCGAGAGAGAGCACTCCGTG 7660
Qy 7321 TGCTGCTCATGTCACTCTCTGAGACGGGGCTTAATTAATCTCTTGTATGCTCCGAAG 7380
Db 7661 TGCTGCTCATGTCACTCTCTGAGACGGGGCTTAATTAATCTCTTGTATGCTCCGAAG 7720
Qy 7381 GAGAAATTACGATTAACCCCTTGAAGCACTCCCTGTGCGATATACAAAGAGGTAC 7440
Db 7721 GAGAAATTACGATTAACCCCTTGAAGCACTCCCTGTGCGATATACAAAGAGGTAC 7780
Qy 7441 TGTACCAACAAGAGCGCTCACTAAGAGGCTTAAGAGTAACTTTGTATGATGACAA 7500
Db 7781 TGTACCAACAAGAGCGCTCACTAAGAGGCTTAAGAGTAACTTTGTATGATGACAA 7840
Qy 7501 GTGCTGACTCTTACTAGACTCACTCACTTTAAAGACATTAAGCTAGCGCTTCAAGTTC 7560
Db 7841 GGCTGAGAGCTCATATGACTCAGTCTTGAAGAGATTAAGCTAGCGCTTCAAGTTC 7900
Qy 7561 ACCGCAAGGCTCTCACTTGAAGAGGCTTGCAGTTAACCCCAACCCATCTGCAAG 7620
Db 7901 ACCGCAAGGCTCTCACTTGAAGAGGCTTGCAGTTAACCCCAACCCATCTGCAAG 7960
Qy 7621 TCTAATATAGGTTTGGGCTTAAGAGGTCGCAAGCTTGTCCGAGAGGCGCTTAAACAC 7680
Db 7961 TCTAATATAGGTTTGGGCTTAAGAGGTCGCAAGCTTGTCCGAGAGGCGCTTAAACAC 8020
Qy 7681 ATCAAGTCCGTGGAAGAGCTCTTGAAGACTCAGAAACACCAATTCACACACATTT 7740
Db 8021 ATCAAGTCCGTGGAAGAGCTCTTGAAGAGACTCAGAAACACCAATTCCTCAACACATC 8080
Qy 7741 ATGGCCAAAATGAGGTGTTTGGCGTGAACCCCAAGGGGGGCAAGAAAGAGCTGCG 7800
Db 8081 ATGGCCAAAATGAGGTGTTTGGCGTGAACCCCAAGGGGGGTAAGAAAGAGCTGCG 8140
Qy 7801 CTATCGTTTACCTGACCTCGGCGTCAAGGCTGCGAGAGATGAGCTCTTATGACATT 7860
Db 8141 CTATCGTTTACCTGACCTCGGCGTCAAGGCTGCGAGAGATGAGCTCTTATGATATC 8200
Qy 7861 ACACAAAAATCTCTCAGGCGGTGATGGGGCTTCTTATGATTCAGATTTCCCGCGCT 7920
Db 8201 ACACAAAAATCTCTCAGGCGGTGATGGGGCTTCTTATGATTCAGATTTCCCGCGCT 8260
Qy 7921 CAGCGGTGAGAGTTTCTTGAAGAGATGGCGGAAAGAGAGCCCTATGGGTTTTCG 7980
Db 8261 CAGCGGTGAGAGTTTCTTGAAGAGATGGCGGAAAGAGAGCCCTATGGGTTTTCG 8320

Qy	661	TTCCACGCTCCCCGGGGTGTGCTCCGGTCCGGAGGAAAGTGGGGGAAATGCACTCAAGTGTGGATA	720
Dp	1001	CTCCACGCTCCCCGGGGTGTGCTCCGGTCCGGAGGAAAGTGGGGGAAATGCACTCCGGTGTGGATA	1060
Qy	721	CCGGTCTCCAGGAATGTGGCCGTGCAGCGGCGCCGGCGCCCTCAACGACAGGGCTTTCGGACG	780
Dp	1061	CCGGTCTCCAGGAATGTGGCCGTGCAGAGCCCGGGCGCCCTCAACGACAGGGCTTTCGGACG	1120
Qy	781	CACATTCACATATGTGTGTGATGTCCGGCCACGCTCTGCTCTGCCCCCTTACGTGTGGGGAACTTC	840
Dp	1121	CACATTCGACATGTGTGTGTGATGTCCGGCCACGCTCTGCTCTGCCCCCTTACGTGTGGGGAACTTC	1180
Qy	841	TGCGGTGGGGGTGATGCTGCAGCCCAATATGTTCAATTGCTGCAGCCGACACCACTGTGTTT	900
Dp	1181	TGCGGTGGGGGTGATGCTGTTCGACGCCAAGATGTTCATTGTCTGCACACAGCACCACTGTGTTT	1240
Qy	901	GTCCAGAGCTGCATATGCTCCATCTCACTCCCTGCTACATCACTGTGACACCGCATGGACATGG	960
Dp	1241	GTGCAGAGCTGCATATGCTCCATCTCACTCCCTGCTACATCACTGTGACACCGCATGGAGCTGG	1300
Qy	961	GACATGATGATGAAGTGTGTGGCCCAACGGCTACAGATCTTGGGGGTACGGCATGCGGTGTC	1020
Dp	1301	GACATGATGATGAAGTGTGTGTGGCCCAACGGCTACAGATCTTGGGGGTACGGCATGCGGTGTC	1360
Qy	1021	CCCGAGGTCAATTATGACATCATTTAGCGGGGCTCATTTGGGGCGTCAATGTTTCGGCTTGGCC	1080
Dp	1361	CCCGAGGTCAATCATGACATCATTTAGCGGGGCTCATTTGGGGCGTCAATGTTTCGGCTTAGCC	1420
Qy	1081	TACTTCTCTTATGACAGGAGCGGTGGCGCAAAATGCTGTGTCAATCCCTTCTGTGGCCGGCGGG	1140
Dp	1421	TACTTCTCTTATGACAGGAGCGGTGGCGCAAAATGCTGTGTCAATCTTCTGTGGCGCGCGGG	1480
Qy	1141	GTGACGCGCGCACCCACTACTGTTGGGGGTTCTGCGCGGACAGACCAACCGGCGGCTCAAC	1200
Dp	1481	GTGACGCGCGCAAAACCCATACCGTGTGGGGTTCTACCGGGCATTAAGCCAGGACCTCTACCC	1540
Qy	1201	AGCTTATTTGACATGGGCCCCAGGCAAGAAATCCAGCTGTGTTAACACCAATGAGCACTGG	1260
Dp	1541	GGCATGTGTTCTCCCTGTGTGTCACAGGCGAAGAAATCCAGCTCATCAACCAATAGGCAGTGGG	1600
Qy	1261	CACATTCACACCGCACCGCCCTGACATGCGAATGATCTCTTGACACACCGCTTATGGCGCT	1320
Dp	1601	CACATTCACACCGCACCGCCCTGAACTGCAATGATCTTTCGACACCGGCTTCTCGCGCTCA	1660
Qy	1321	CTGTTTCTACACCCACAGCTTCAACTCTGCAGAGATGTCCGGAACGATGTCCGCTGCGCG	1380
Dp	1661	CTGTTTCTACACCCACAGCTTCAACTCTGCAGAGATGTCCGGAACGATGTCCGCTGCGCGC	1720
Qy	1381	AGTATTCAGGCGCTTCCGGGTGGAGATGGGGCGCTTGGCAATATGAGATATATGTCAACCAAT	1440
Dp	1721	AGTATTCAGGCGCTTCCGGGTGGAGATGGGGCGCTTGAACAATATGAGAACAAATGTCAACCAAT	1780
Qy	1441	CCAGAGATATGAGACCTTATTTGCTGCGCACTACCAACCAAGGCAAGTGTGGCGTGTCTCC	1500
Dp	1781	CCAGAGGATATGAGACCTGATTTGCTGCGCACTACCAACCAAGACAGTGTGTGTACTCTCC	1840
Qy	1501	GCGAAGACTGTGTGTGGCCCCAGTGTACTGTTTACCCCCAGCGCCAGTGTATGTGGGACG	1560
Dp	1841	GCGAAGCTCTGTGTGTGGCCCCAGTGTACTGTTTACCCCCAGCGCCAGTGTATGTGGGTACG	1900
Qy	1561	ACCGACAGGCTTGGAGCGGCCACTTACACGTGGGGGGGAAATGAGACAGATGTCTTCCCTA	1620
Dp	1901	ACCGATAGACTTGGAGCGGCCACTTACACGTGGGGGGGAAATGAGACAGATGTCTTCCCTA	1960
Qy	1621	TTGAACAGCACTCGAACCGCTGGGGTTCATAGTTCGGCTGCACCTGTGATCAACTTCTT	1680
Dp	1961	TTGAACAGGACTCGAACCGAGGGGTTCATAGTTCGGCTGCACCTGTGATGAATCACTCACT	2020
Qy	1681	GGCTTACACCAAGACTTGGGGCGCACCACTCTGCGTACTAGACTGACTTCAAGCGCACG	1740
Dp	2021	GGCTTACACCAAGACTTGGGGCGCACCACTCTGCGCTTAAAGCTGACTTCAATGCGCACG	2080

QY	1741	ACGGACCTGGTGTGGCCCAACGGACCTGTTTTAGGAAACATCCGTGATACCACTTACCTCAAA	1800
Db	2081	ATGGACTGTGTGTGCCCAACGGACCTGTTTTAGGAAACATCTGTATACCACTTACATCAA	2140
QY	1801	TGCGGCTCTGGGCCCTGTGGCTCAACGCCAAGTGCCTGATCGACTACCCCTTACAGGCTCTGG	1860
Db	2141	TGTGGCTCTGGGCTCCCTGGGCTCAACGCCAAGTGCCTGATCGACTACCCCTTACAGGCTCTGG	2200
QY	1861	CATTACCCCTGACAGTTATATCTATACCATCTTCAAAATTAAGATGTATGTGGAGGGGTT	1920
Db	2201	CATTACCCCTGACAGTTATATCTATACCATCTTCAAAATTAAGATGTATGTGGAGGGGTT	2260
QY	1921	GAGCAACGGCTCAACGGCTGCGATGCAATTTCACTCGTGGGGATCGTTGCCAATTGGAGAC	1980
Db	2261	GAGCAACGGCTCAACGGCTGCGATGCAATTTCACTCGTGGGGATCGTTGCCAATTGGAGAC	2320
QY	1981	AGAGACAGAAATGCAACTGTCTCTCTTTGTGGACTCCACCAACGGAAATGGGCAATTTTACTT	2040
Db	2321	AGAGACAGAAATGCAACTGTCTCTCTTTGTGGACTCCACCAACGGAAATGGGCAATTTTACTT	2380
QY	2041	TGCTCTTACTGGGACCTGGCCGGCTGTGCGACTGTGCTTCTCCAGCTCCACCAAAACATC	2100
Db	2381	TGCACTTACTCGGACCTGGCCGGCTGTGCGACTGTGCTTCTCCAGCTCCACCAAAACATC	2440
QY	2101	GTGACGTACAAATTCATGTATGTGGCTTATGACCTGCGCTCACAAAATATCATGTCCGATGG	2160
Db	2441	GTGACGTGTCAATTCATGTATGTGGCTTATGACCTGCGCTCACAAAATATCATGTCCGATGG	2500
QY	2161	GAGTGGGTATATCTCTTATTCGCGCTCTTAAGCGGAGCGCAGGGTTTGGCGCTGTATGG	2220
Db	2501	GAGTGGGTATATCTCTTATTCGCGCTCTTAAGCGGAGCGCAGGGTTTGGCGCTGTATGG	2560
QY	2221	ATGCTCATCTTGTGGGCGCAGGCCGAGAGCAGCATAGAGAAAGCTGTGATCTTGGCACGCT	2280
Db	2561	ATGCTCATCTTGTGGGCGCAGGCCGAGAGCAGCATAGAGAAAGCTGTGATCTTGGCACGCT	2620
QY	2281	GGGAGCGCAGCTAGCTGCAATGGCTTCCNATATTTTGTGATCTTTTTCGTGGCGCTTGG	2340
Db	2621	GGGAGCGCAGCTAGCTGCAATGGCTTCCNATATTTTGTGATCTTTTTCGTGGCGCTTGG	2680
QY	2341	TACATCAAGGGTCGGGTAGTCCCTCTTACGTACCTATTCCTCACTGGCGCTGTGGCTCTT	2400
Db	2681	TACATCAAGGGTCGGGTAGTCCCTCTTACGTACCTATTCCTCACTGGCGCTGTGGCTCTT	2740
QY	2401	AGCCTACTGCTCTAGCATTTGCCCAACAGGCTTATGTATGACGATCTGTGCATGGC	2460
Db	2741	GGCCTACTGCTCTAGCATTTGCCCAACAGGCTTATGTATGACGATCTGTGCATGGT	2800
QY	2461	CAGATAGAGACGGCTGTCTGTGTATAGATCACTCTCTTATCTCAACCCCGGGTATAG	2520
Db	2801	CAGATAGAGACGGCTGTGTGTATAGATCACTCTCTTATCTCAACCCCGGGTATAG	2860
QY	2521	ACCCCTTCAGCCCGGTTTTGTGTGGTGTGTGCTATCTTGTGACCCTGGGGGAAAGCTATG	2580
Db	2861	ACCCCTTCAGCCCGGTTTTGTGTGGTGTGTGCTATCTTGTGACCCTGGGGGAAAGCTATG	2920
QY	2581	GTTCAGAGATGGGCAACCACTATGCAAGTGCGGGTGGCCGTATGCGCATATATGGGCC	2640
Db	2921	GTTCAGAGATGGGCAACCACTATGCAAGTGCGGGTGGCCGTATATGGAAATATATGGGCC	2980
QY	2641	GTTCGCAATATTTCAACCAAGTGTGTGTGTGACATAAACAAGTGTGCTTGTGGCGGTGTT	2700
Db	2981	GTTCGCAATATTTCTGCCCGGGTGTGTGTGTGACATAAACAAGTGTGCTTGTGGCGGTGTT	3040
QY	2701	GGGCGTGTACTCTCTTAAAGGTGTGTTGACGCGCGTGCCTGATCTTCTGTCAGGGCTCAC	2760
Db	3041	GGGCGTGTACTCTCTTAAAGGTGTGTTGACGCGGTGCCTGATCTTCTGTCAGGGCTCAC	3100
QY	2761	GCTTACTATGAGATGTGCAACATGGGAGGCACTGCGCGGGGGCAGGTAGCTCCAGATG	2820
Db	3101	GCTTACTATGAGATGTGCAACATGGGAGGCACTGCGCGGGGGTATGATGTCCAGATG	3160
QY	2821	GGGCTACTAGCCCTTGGCAGGTGTGACTGTGGCACTTATATGACCACTTACCCCTATG	2880

```
Db 3161 GTGCTACTAGCCCTTGGCAGGTGGAAGGACCTTATGATGACCACTTACCCCTATG 3220
Qy 2881 TCGGATTTGGGCTGCTAGTGGCTTGGGACCTGGCGGTCCGCTTGAAGCTTATCATCTTC 2940
Db 3221 TCGGATTTGGGCTGCTAGTGGCTTGGGACCTGGCGGTCCGCTTGAAGCTTATCATCTTC 3280
Qy 2941 AGTCCGATGGAAGAAAGTCACTTGTCTGGGAGCGGAGACAGCTGCTTGTGGGAGCACTT 3000
Db 3281 AGTCCGATGGAAGAAAGTCACTTGTCTGGGAGCGGAGACAGCTGCTTGTGGGAGATTC 3340
Qy 3001 TTACACGGAAGCTTCCGCTGTCGCGCCGACCTTGTGGAGGCTCTCTTGGCCGACGTAT 3060
Db 3341 TTACACGGAAGCTTCCGCTGTCGCGCCGACCTTGTGGAGGCTCTCTTGGCCGACGTAT 3400
Qy 3061 GGTATTAACCTCCAAAGGGGTGGAATCTTCTGCGCCGCTCACTGCTTACGCGCAGACACA 3120
Db 3401 GGTATTAACCTCCAAAGGGGTGGAATCTTCTGCGCCGCTCACTGCTTACGCGCAGACACA 3460
Qy 3121 CGTGGCCTTTTGGGACCATAGTGTGAGCATGACGGGGGCGGCAAGACAGAGAGGCT 3180
Db 3461 CGGCGCTTTTGGGACCATAGTGTGAGCATGACGGGGGCGGCAAGACAGAGAGGCT 3520
Qy 3181 GGGGAAATTCAGGTCCTGTGTCACAGTCACTCAGTCTTCTCGGAACATCATCTCGGG 3240
Db 3521 GGGGAAATTCAGGTCCTGTGTCACAGTCACTCAGTCTTCTCGGAACATCATCTCGGG 3580
Qy 3241 GTTTTGTGACCTGCTTACCATGAGCTGAGCAACAAGCTTGCGCGGCTACCGGGGTCCG 3300
Db 3581 GTCTTATGACCTGTCTTACCATGAGCTGAGCAACAAGCTTACCGGCTTACCGGGGTCCG 3640
Qy 3301 GTACACGAGATGATCTCAGTGTGAGGGGAGCTTATGATGGGTGGCCGACCCGCTGGG 3360
Db 3641 GTACACGAGATGATCTCAGTGTGAGGGGAGCTTATGATGGGTGGCCGACCCGCTGGG 3700
Qy 3361 ACTAAATCTTTGGAGCCGTGACAGTGTGAGCGGTGACCTGTATCCTGATCAGCGGAGAC 3420
Db 3701 ACCAAATCTTTGGAGCCGTGACAGTGTGAGCGGTGACCTGTATCCTGATCAGCGGAGAC 3760
Qy 3421 GCTGATGTCACTCCCGGCTCGAAGACCGGGGAGCAAAAGGGAGGCTTACTCTCCCGAGA 3480
Db 3761 GCTGATGTCACTCCCGGCTCGAAGACCGGGGAGCAAAAGGGAGGCTTACTCTCCCGAGA 3820
Qy 3481 CCTCTTTCACCTTGAAGGGGCTCTCAGAGAGCCGGGTGCTATGCCCAGGGGCAAGCT 3540
Db 3821 CCTCTTTCACCTTGAAGGGGCTCTCAGAGAGCCGGGTGCTATGCCCAGAGGCAAGCT 3880
Qy 3541 GTGGAAGTCTTCCGGGACGCTGTGTGCTCTGGGGGCTGTGCTAAGTCCATAGAATTCATC 3600
Db 3881 GTGGAAGTCTTCCGGGACGCTGTGTGCTCTGGGGGCTGTGCTAAGTCCATAGAATTCATC 3940
Qy 3601 CCCGTTGAGACATCGACATGTCACGCGGTCCGCCACCTTTTAGTGAACAAGACACA 3660
Db 3941 CCCGTTGAGACATCGACATGTCACGCGGTCCGCCACCTTTTAGTGAACAAGACACA 4000
Qy 3661 CCGTGTGTGCGCCAGACTATCAGGTGCGGCTTATGATGCCCCGAGTGGCAGTGAAG 3720
Db 4001 CCGTGTGTGCGCCAGACTATCAGGTGCGGCTTATGATGCCCCGAGTGGTGAAG 4060
Qy 3721 AGCACCAGAAAGTCTGTGCTATGCTGCTCAGGGGATTAAGTGTGCTAAGTAAATCCC 3780
Db 4061 AGCACCAGAAAGTCTGTGCTATGCTGCTCAGGGGATTAAGTGTGCTAAGTAAATCCC 4120
Qy 3781 TCAAGTGGCTGCCACCTCGGGGTTTGGGGCTTACTTGTCTTAAAGCAGATGGCATTC 3840
Db 4121 TCAAGTGGCTGCCACCTCGGGGTTTGGGGCTTACTTGTCTTAAAGCAGATGGCATTC 4180
Qy 3841 AACATTAGGACTGGAATCAGAGCTGTGACGACGCGGGGCGCCATCAGTACTCAACTAT 3900
Db 4181 AACATTAGGACTGGAATCAGAGCTGTGACGACGCGGGGCGCCATCAGTACTCAACTAT 4240
Qy 3901 GGCAAATTCCTCGCCGATGGGGGCTGTGCGGGCGGCGCTTACGACATCATATGTGAT 3960
Db 4241 GGCAAATTCCTCGCCGATGGGGGCTGTGCGGGCGGCGCTTACGACATCATATGTGAT 4300
Qy 3961 GAATGCCATGCGGTGACCTTACACCAATCCTTGGCATTCGGAACAGTCTTATTCAGCA 4020
Db 4301 GAATGCCATGCGGTGACCTTACACCAATCCTTGGCATTCGGAACAGTCTTATTCAGCA 4360
Qy 4021 GAGACAGCTGGGGTACAGACTTAACTGTGCTGGCTTACAGCTACGCGCCCTGGGTCAGTACA 4080
Db 4361 GAGACAGCTGGGGTACAGACTTAACTGTGCTGGCTTACAGCTACGCGCCCTGGGTCAGTACA 4420
Qy 4081 ACCCCGCAACCCCAATAGAGAGTGGCCCTTGGGAGAGGCGGAGATCCCTTGTAT 4140
Db 4421 ACCCCGCAACCCCAATAGAGAGTGGCCCTTGGGAGAGGCGGAGATCCCTTGTAT 4480
Qy 4141 GGGAGGGCGATTCGCCCTGTATCATCAAGGGAGGAAGATCTGATCTTGTGCATTTCA 4200
Db 4481 GGGAGGGCGATTCGCCCTGTATCATCAAGGGAGGAAGATCTGATCTTGTGCATTTCA 4540
Qy 4201 AAGAAAAAGTGTACGAGCTCGCGCGGCGCTTCGGGGTATGGGCTTGAATCAGTGGCA 4260
Db 4541 AAGAAAAAGTGTACGAGCTCGCGCGGCGCTTCGGGGTATGGGCTTGAACGAGTGGCA 4600
Qy 4261 TACTACAGAGGTTTGAACGTCTTCCTGTAATCAACTCAGGAGACGTATGTGTGCTGCC 4320
Db 4601 TACTACAGAGGTTTGAACGTCTTCCTGTAATCAACTCAGGAGACGTATGTGTGCTGCC 4660
Qy 4321 ACCGAGCCCTCATGACAGGATTAATCTGGGAGCTTTGATCTCGGTGATCTGCAAGTGA 4380
Db 4661 ACCGAGCCCTCATGACAGGATTAATCTGGGAGCTTTGATCTCGGTGATCTGCAAGTGA 4720
Qy 4381 GCGGTCACTCAAGTTTGAAGCTTCAAGTTTGAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 4440
Db 4721 GCGGTCACTCAAGTTTGAAGCTTCAAGTTTGAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 4780
Qy 4441 GTCCCTTCAAGCGCTGTCTTCAAGTACGAGCGCGCGGCTTCAAGCTTCAAGCTTCAAG 4500
Db 4781 GTCCCTTCAAGCGCTGTCTTCAAGTACGAGCGCGCGGCTTCAAGCTTCAAGCTTCAAG 4840
Qy 4501 GGCATTTATAGATATGTTTCCACTGATGAGAGCGGCTTCAAGGATTTTGAACAGTGTG 4560
Db 4841 GGCATTTATAGATATGTTTCCACTGATGAGAGCGGCTTCAAGGATTTTGAACAGTGTG 4900
Qy 4561 CTCTGTAGAGTCAACACGAGAGGCGCGATGATGAGCTCAACCATCGAGAGACACC 4620
Db 4901 CTCTGTAGAGTCAACACGAGAGGCGCGATGATGAGCTCAACCATCGAGAGACACC 4960
Qy 4621 GTCAGGCTCAGAGCATATTTCAACACACCTGTGCTGTGTCGCAAGCACTTTAG 4680
Db 4961 GTCAGGCTCAGAGCATATTTCAACACACCTGTGCTGTGTCGCAAGCACTTTAG 5020
Qy 4681 TTTTGGAGGCACTTTTCAACCGGCTTCAACACATAGATGCCCACTTCTTTCCAAACA 4740
Db 5021 TTTTGGAGGCACTTTTCAACCGGCTTCAACACATAGATGCCCACTTCTTTCCAAACA 5080
Qy 4741 AAGCAATCGGGGAGAAATTTGCGATCTTAAAGAGCTTACAGAGCTACAGTGTGCGTAGG 4800
Db 5081 AAGCAATCGGGGAGAAATTTGCGATCTTAAAGAGCTTACAGAGCTACAGTGTGCGTAGG 5140
Qy 4801 GCCAAAGCCCCCCCCGCTGTGGAAGCTCATGTGAGAGTGTGACTCGACTCAAGCCC 4860
Db 5141 GCCAAAGCCCCCCCCGCTGTGGAAGCTCATGTGAGAGTGTGACTCGACTCAAGCCC 5200
Qy 4861 ACACTGTGGGCCCCCAGACTCTCTGTGACCGTGTGGGCTGTGTACCAAGAGGTAC 4920
Db 5201 ACACTGTGGGCCCCCAGACTCTCTGTGACCGTGTGGGCTGTGTACCAAGAGGTAC 5260
Qy 4921 CTCACACATCCCGTGAAGAAATACATGCGCACTGTGATGCAAGCCGACTTGAAGTATG 4980
Db 5261 CTCACACATCCCGTGAAGAAATACATGCGCACTGTGATGCAAGCCGACTTGAAGTATG 5320
Qy 4981 ACCAGACATGGGCTTTGGAGGGGAGATCTTGGCGGCGGTGCGCGGTATGCTGGCG 5040
Db 5321 ACCAGACATGGGCTTTGGAGGGGAGATCTTGGCGGCGGTGCGCGGTATGCTGGCG 5380
```

OY	5041	ACCGGCTGTGTTGATCATATGCGCGCTTGGACATTTAAACAAGGAGCGCTGTTGGCGG	5100
DB	5381	ACCGGCTGTGTTGATCATATGCGCGCTTGGACATTTAAACAAGGAGCGCTGTTGGACCG	5440
OY	5101	GACAAAGAGGCTCTCTATGAGGCTTTTGATGAGATGAGAGATGCTCTAGGAGCGCT	5160
DB	5441	GACAAAGAGGCTCTCTATGAGGCTTTTGATGAGATGAGAGATGCTCTAGAGCGGCT	5500
OY	5181	CTCATTTGAAAGAGGGCAGCGGATAGCCGAGATGCTGAAGTCCAAAGATCCAAAGCTTATTG	5220
DB	5501	CTCATTTGAAAGAGGGCAGCGGATAGCCGAGATGCTGAAGTCCAAAGATCCAAAGCTTATTG	5560
OY	5221	CAGCAAGGCTTCCAAACAAGCTCAAGACATTAACAACCACTGTGCAGAGCTTACATGCCCCAAG	5280
DB	5561	CAGCAAGGCTTCCAAACAAGCTCAAGACATTAACAACCCGCTGTGCAGAGCTTCTTGCCCCAAG	5620
OY	5281	GTAGAACAATTCTGGGCCCAACACATGTGGAATTCACTTAGCGGCATCCAATACCTTGCA	5340
DB	5621	GTAGAGCAATTCTGGGCCCAACACATGTGGAATTCACTTAGCGGCATCCAATACCTTGCA	5680
OY	5341	GGACATATTAACATCGCCAGCGGAAACCTCTGACATTCCTCAAGATGGGCTTCAGTCCGCC	5400
DB	5681	GGACATATTAACATCGCCAGCGGAAACCTCTGACATTCCTCAAGATGGGCAATTCAGTCCGCC	5740
OY	5401	CTCACAGATCGCTGTCAACAAGCACACATATCCCTTCAACATTTTGGGGGCTGTGCTA	5460
DB	5741	CTCACAGATCGCTGTGTCAACAAGCACATATCCCTTCAACATTTTGGGGGCTGTGCTA	5800
OY	5461	GCATCCCAATTGTCACCAACCGCGGGGGGCACCTGGCTTGTGTGATGTGCCTAGTGGGA	5520
DB	5801	GCATCCCAAAATGGCCCTCCCGCGGGGGCTACCGGCTTGTGTGATGTGCCTAGTGGGG	5860
OY	5521	GCTGCCGATAGGACGATATAGGCTTAGTGAAGGCTAGTGAACATCTGTGCAGGGTATAGT	5580
DB	5861	GCTGCCGATAGGACGATATAGGCTTAGGCTTAGGCTGTGTGACATCTGTGCAGGGTATAGT	5920
OY	5581	GGGGGCATTTTGGGGGCTCTGTGTGCATTCAGATCATGTCTGTGCAGAAAGCCCTCCATG	5640
DB	5921	GGGGGCATTTTGGGGGCTCTGTGTGCATTCAGATCATGTCTGTGCAGAAAGCCCTCCATG	5980
OY	5641	GAGGATGTCTCAATTGCTGTGCTGTGGGAATTCGTCTCCGGGTGCTGGTATGTGGAGTCT	5700
DB	5981	GAGGATGTGTCAACTCTGTGTGTGGGAATTCGTCTCCGGGTGCTGGTATGTGGAGTCT	6040
OY	5701	ATCTCGCGGGCCATCTCGCGCGACGACGCTGGGACCGGGGGAAGCGCGCTCAATGTGATG	5760
DB	6041	ATCTCGCGGGCCATCTCTGCGCGACGACGCTGGGACCGGGGGAAGCGCTGTCAATGTGATG	6100
OY	5761	AATAGACTCATTTGCTTTTGCTTCCAGAGGAAATCAGCTGCGCCCACTCACTACGTGACG	5820
DB	6101	AATAGACTCATTTGCTTTTGCTTCCAGAGGAAACCACTGCGCCCACTCACTACGTGACG	6160
OY	5821	GAGTGGGATGCGTGTGCAGCGGTGACCCAACTACTTGGCTCCCTTACCACTAACAGGCTG	5880
DB	6161	GAGTGGGATGCGTGTGCAGCGGTGACCCAACTACTTGGCTCCCTTACCACTAACAGGCTG	6220
OY	5881	CTCAGAAAGCTTCACAACTGGAATTACTGAGAGCTGCCCCCATCCCATGCGCGGCTGTGG	5940
DB	6221	CTCAGAAAGCTTCACAACTGGAATTACTGAGAGCTGCCCCCATCCCATGCGCGGCTGTGG	6280
OY	5941	CTCCCGCATGTGTGGGACTGGGTTTGGACATCTTAAGAATTTAAATTTGGCTGACC	6000
DB	6281	CTCCCGCATGTGTGGGATTTGGGTTTGGACATCTTAAGAAGCTTTAAATTCGCTGACC	6340
OY	6001	TCCAATTTATTTCCCAAAGATGCCCCGCTCCCTTTGTCTCTGTCAAAAAGGGGTACAG	6060
DB	6341	TCCAATTTGTTCCCAAAGATGCTGTGTCTCCCTTTATCTCTGTCAAAAAGGGGTACAG	6400
OY	6061	GGCGTGTGGCGCGGCACTGTGCATATGACACACGCTGTCTTGGGGCGCAATATCTCT	6120
DB	6401	GGCGTGTGGCGCGGCACTGTGTATATGACACACGCTGTCTTGGGGCGCAATATCTCT	6460

QY	6121	GGCAATGTCGGCTTGGGCTCCATAGAAATACCGGGGCGTAAAGCTCCATGAATATCTGG	6180
Db	6461	GGCAATGTCGGCTTGGGCTCCATAGAAATTAAGGGGCCAACAACCTCATAAGATATCTGG	6520
QY	6181	CAGGGGACCTTTCCTATCAATTGTTTAAACGAGGGGCCAGTGCCTGCCAAACCCGCGCA	6240
Db	6521	CAGGGGACCTTTCCTATCAATTGTTTAAACGAGGGGCCAGTGCCTGCCAAACCCGCGCA	6580
QY	6241	AACCTTAAAGTTCGCCATCTGGAAGGGTGGCGGCTCAGAGTAACGGAGAGTGAACGACAC	6300
Db	6581	AACCTTAAAGTTCGCCATCTGGAAGGGTGGCGGCTCAGAGTAACGGAGAGTGAACGACAC	6640
QY	6301	GGGTCAATCACTACATPAACAGACCTCACACTGATTAACCTTGAAGTCCCTGCGCACTA	6360
Db	6641	GGGTCAATCACTACATPAACAGACCTTACACTGATTAACCTTGAAGTTCCTGCGCACTA	6700
QY	6361	CCCTCTCCGAGTTCTTTCTTGGGTGACGGAATGACATCCATAGTTTGCCTCCACA	6420
Db	6701	CCCTCTCCGAGTTCTTTCTTGGGTGACGGAATGACATCCATAGTTTGCCTCCACA	6760
QY	6421	CCGAAGCCGTTTTCCGGGAGAGAGTCGTCGTCGCTTGGGCGCTAATTTCAATTGTGTC	6480
Db	6761	CCGAAGCCGTTTTCCGGGAGAGAGTCGTCGTCGCTTGGGCGCTAATTTCAATTGTGTC	6820
QY	6481	GGGTCCCACTCTCTTGGACCCCTGAAACCCGACACAGACGTATTTGATGTCCATGCTACA	6540
Db	6821	GGGTCCCACTCTCTTGGACCCCTGAAACCCGACACAGACGTATTTGATGTCCATGCTACA	6880
QY	6541	GATCATTTCAATCAACGGCGGAACTGACAGCGCGGCTTTAGCGCGGGGTCAACCCCA	6600
Db	6881	GATCATTTCAATCAACGGCGGAACTGACAGCGCGGCTTTAGCGCGGGGTCAACCCCA	6940
QY	6601	TCCGAGGCAACTCCTCGCGAGGCAAGCTATCGACACCATGCTGCGAGCCACTGCAAC	6660
Db	6941	TCCGAGGCAACTCCTCGCGAGGCAAGCTATCGACACCATGCTGCGAGCCACTGCAAC	7000
QY	6661	ACCCACGGCAAGCCTATGATGTGACATGTGATGCTAACCTGTTCATGCGGGGCGAT	6720
Db	7001	ACCCACGGCAAGCCTATGATGTGACATGTGATGCTAACCTGTTCATGCGGGGCGAT	7060
QY	6721	GTGATCGGATAGAGTCTGGGTCCAAAGTGTGCTGTCGACTGCTCCGACCCCAATGCTC	6780
Db	7061	GTGATCGGATAGAGTCTGGGTCCAAAGTGTGCTGTCGACTGCTCCGACCCCAATGCTC	7120
QY	6781	GAGAGAAAGACGACCTTGAGCTTCGATACCATCAGATTAACATGCTCCCAAGAGAGG	6840
Db	7121	GAGAGAAAGACGACCTTGAGCTTCGATACCATCAGATTAACATGCTCCCAAGAGAGG	7180
QY	6841	TTTCCACCAAGCTTAAACGGGCTGGGACAGGCTGATTAACAACCAACGGCTTGTGGAATCG	6900
Db	7181	TTTCCACCAAGCTTAAACGGGCTGGGACAGGCTGATTAACAACCAACGGCTTGTGGAATCG	7240
QY	6901	TGGAAAAGGCGAGATTACCAACCGGCGCACTGTCGGGGCTGTGCTCTCCCTCCTAGG	6960
Db	7241	TGGAAAAGGCGAGATTACCAACCGGCGCACTGTCGGGGCTGTGCTCTCCCTCCTAGG	7300
QY	6961	AAAAACCCGACGCTCTCCCAAGAGAGCGCGGACAGTGGGTCTGAGTGAAGCTTCATA	7020
Db	7301	AAAAACCCGACGCTCTCCCAAGAGAGCGCGGACAGTGGGTCTGAGTGAAGCTTCATA	7360
QY	7021	GGAGATGCGCTTCAACAGCTGAGCATTAAAGTCTTTGGCCAGGCCCGCCCAAGGGCGAT	7080
Db	7361	GGAGATGCGCTTCAACAGCTGAGCATTAAAGTCTTTGGCCAGGCCCGCCCAAGGGCGAT	7420
QY	7081	TCAGGCTTTTCAACGGGGGCGGGCGCTGCGATTTCCGGCACTGACACGCTCTTGATAG	7140
Db	7421	TCAGGCTTTTCAACGGGGGCGGGCGCGAGCTGATTTCCGGCACTGCGAGCGCCCGCATAG	7480
QY	7141	TTGGGCTTTTGGGAGACAGGTTTCAATCTCTTCATGCCCCCTCTCGAAGGGGAGCTTGG	7200
Db	7481	TTGGGCTTTTGGGAGACAGGTTTCAATCTCTTCATGCCCCCTCTCGAAGGGGAGCTTGG	7540
QY	7201	GATCCAGACTGAGCTGAGCTGAGCTGATAGGCCCAACCCCGCCCAAGGGGGGAGTGGCA	7260

Db 7541 GATTCAGACTTGGAGCTTGAGAGGTGAGACTTCAACTCTCCCCCGAGGGGGGGTGTGA 7600
Qy 7261 GCTCCCGGCTCGAGACTGGGGTCTGTGTACTTGTCTCCGAGAGAGCACTCGTGTG 7320
Db 7601 ACCCGGGCTCAGGCTCGGGGTCTTGTCTACTTGTCTCCGAGAGAGCACTCGTGTG 7660
Qy 7321 TGTGCTCATCTGATCTCTGTGACCGGGGCTTAACTAATCTCTGTGTAGTCCGAAAG 7380
Db 7661 TGTGCTCATCTGATCTCTGTGACCGGGGCTTAACTAATCTCTGTGTAGTCCGAAAG 7720
Qy 7381 GAGAGTTACCGATTAACTCCCTTGAGCACTCCCTGTGGATATCAACAAGGTGTAC 7440
Db 7721 GAAAGTTGCGAATGGAGCTTGAGCACTCCCTGTGGATATCAACAAGGTGTAC 7780
Qy 7441 TGTACACAAACAAGAGCGCTCACTAAGGGCTTAAAGGTAACTTTGTAGATGCA 7500
Db 7781 TGTACACAAACAAGAGCGCTCACTAAGGGCTTAAAGGTAACTTTGTAGATGCA 7840
Qy 7501 GTGCTGACTCTACTACGACTCAGTCTTAAAGAGACTTAAGGAGGCTCGCAAGTCT 7560
Db 7841 GCGCTGAGCGCTCATATGACTCAGTCTTAAAGAGACTTAAGGAGGCTCGCAAGTCT 7900
Qy 7561 ACCGCAAGGCTCTCAACATGAGAGAGCTTGCAGTTAACTCCCACTTGTGCAAG 7620
Db 7901 ACCGCAAGGCTCTCAACATGAGAGAGCTTGCAGTTAACTCCCACTTGTGCAAG 7960
Qy 7621 TCTAAATATGGGTTTGGGGGTAGAGAGTCCGAGCTTGTCCGGAGGCGCTTAACT 7680
Db 7961 TCTAAATATGGGTTTGGGGGTAGAGAGTCCGAGCTTGTCCGGAGGCGCTTAACT 8020
Qy 7681 ATCAAGTCCGTGAGAGAGACTCTGAGAGAGCTCAAGAAACCAATTTCCACACAT 7740
Db 8021 ATCAAGTCCGTGAGAGAGACTCTGAGAGAGCTCAAGAAACCAATTTCCACACAT 8080
Qy 7741 ATGGCAAAAATAGAGTGTCTGTGAGACCCCAACAAGGGGGCAAGAAAGAGCTGCG 7800
Db 8081 ATGGCAAAAATAGAGTGTCTGTGAGACCCCAACAAGGGGGTAAGAAAGAGCTGCG 8140
Qy 7801 CTATATGTTTAACTCTGAGCTCGGCGCTCAGGGCTTGTGAGAAATGCGCTTATGAT 7860
Db 8141 CTATATGTTTAACTCTGAGCTCGGCGCTCAGGGCTTGTGAGAAATGCGCTTATGAT 8200
Qy 7861 ACACAAAACCTCTGAGCGGTGATGGGGGCTTCTATGATTTCCAGATTTCCCGGCT 7920
Db 8201 ACACAAAACCTCTGAGCGGTGATGGGGGCTTCTATGATTTCCAGATTTCCCGGCT 8260
Qy 7921 CAGCGGGTGAAGTCTTCTTGAAGAAGTGGCGGAAAGAGAGCCCTATGGGTTTTGCG 7980
Db 8261 CAGCGGGTGAAGTCTTCTTGAAGAAGTGGCGGAAAGAGAGCCCTATGGGTTTTGCG 8320
Qy 7981 TATGATATCCGATGCTTGAAGTCAACCGTCACTGAGAGAGACTCAGAGTGAAGTCC 8040
Db 8321 TATGATATCCGATGCTTGAAGTCAACCGTCACTGAGAGAGACTCAGAGTGAAGTCC 8380
Qy 8041 ATATATCGGGCTGCTCTTGGCCGAGAGGCCACACTGSCATACACTGCTAACTAG 8100
Db 8381 ATATATCGGGCTGCTCTTGGCCGAGAGGCCACACTGSCATACACTGCTAACTAG 8440
Qy 8101 AGACTTATCGTGGAGGGCTTATGTTCAACAGCAAGGCCAACTGCGGGTCAAGCGCT 8160
Db 8441 AGACTTATCGTGGAGGGCTTATGTTCAACAGCAAGGCCAACTGCGGGTCAAGCGCT 8500
Qy 8161 TGGCGGCGCAGCGGGTGTCTCAACACTAGATGGGGAAACCACTACATCTACTGTA 8220
Db 8501 TGGCGGCGCAGCGGGTGTCTCAACACTAGATGGGGAAACCACTACATCTACTGTA 8560
Qy 8221 GCTTATGCGGCTTGTAAAGCTGAGGAGATTAATGCGGCCCAATGCTGTGATCGGCGAT 8280
Db 8561 GCTTATGCGGCTTGTAAAGCTGAGGAGATTAATGCGGCCCAATGCTGTGATCGGCGAT 8620
Qy 8281 GACTTGTGTCTCATCTCAGAAAGCCAGGGAGCCGAGAGAGCAGCGGAACTGTAGAGCC 8340

Db 8621 GACTGTGTGTCTCATCTCAGAGAGCCAGGGAGCCGAGAGGAGCAGCGGAACCTGAGAGCC 8680
Qy 8341 TTACAGGAGGCTATGACACAGATTTTGTGCCCCCTCTGTGTAAACCCCCCAAGCCGAGTAT 8400
Db 8681 TTACAGGAGGCTATGACACAGATTTTGTGCCCCCTCTGTGTAAACCCCCCAAGCCGAGTAT 8740
Qy 8401 GATCTGAGGCTGATTAACATCTTGTCTCTCAATATGTGTGTGTGAGCGTGGCCCAAGAGC 8460
Db 8741 GACTGTGAGGCTGATTAACATCTTGTCTCTCAATATGTGTGTGTGAGCGTGGCCCAAGAGC 8800
Qy 8461 CGCGGAGATCTACTGACAGAGACCTTACCACTTCAATGCGCGGGCTGCTGGAA 8520
Db 8801 CGCGGAGATCTACTGACAGAGACCTTACCACTTCAATGCGCGGGCTGCTGGAA 8860
Qy 8521 ACAGTTAGACATCCCTCTGCAATTTATGAGCTGGGAAACATCATCCAGTACGCCCGAAC 8580
Db 8861 ACAGTTAGACATCCCTCTGCAATTTATGAGCTGGGAAACATCATCCAGTACGCCCGAAC 8920
Qy 8581 ATATGGGCTCGCATGTGCTCTGATGACACACTTCTCTCATTTCTCATAGGCTCAAGACAGC 8640
Db 8921 ATATGGGCTCGCATGTGCTCTGATGACACACTTCTCTCATTTCTCATAGGCTCAAGATCT 8980
Qy 8641 CTGGACCAAGACCTCAACTTTGATGATGACGAGCGGTGACTCTCGTGAATCCCTTGAC 8700
Db 8981 CTGGACCAAGACCTCAACTTTGATGATGACGAGCGGTGACTCTCGTGAATCCCTTGAC 9040
Qy 8701 CTCCGAGCTTAAATTAAGAGTTACATGGGCTTGAAGCTTTTCTCTGCAACATACAT 8760
Db 9041 CTCCGAGCTTAAATTAAGAGTTACATGGGCTTGAAGCTTTTCTCTGCAACATACAT 9100
Qy 8761 CCCACGAACCTGACACAGGGTGTCTGAGCCCTGAGAAACTTGGGGGCGCCACCTCAGA 8820
Db 9101 CCCACGAACCTGACACAGGGTGTCTGAGCCCTGAGAAACTTGGGGGCGCCACCTCAGA 9160
Qy 8821 GCGTGAAGAGCCGGGACGTGAGTCAAGGGCTCTCTCATCTCCGTGGGGGAGAGCG 8880
Db 9161 GCGTGAAGAGCCGGGACGTGAGTCAAGGGCTCTCTCATCTCCGTGGGGGAGAGCG 9220
Qy 8881 GCGGTTTGGGTGATATCTCTTCAATTTGGGCGGTGAAGACCAAGCTCAAACTCACTCCA 8940
Db 9221 GCGGTTTGGGTGATATCTCTTCAATTTGGGCGGTGAAGACCAAGCTCAAACTCACTCCA 9280
Qy 8941 TTGCGGGAAGCGGCTCTGATTTATCAAGTGTGTCACGTCGCGCGCGGGGCGG 9000
Db 9281 TTGCGGGAAGCGGCTCTGATTTATCAAGTGTGTCACGTCGCGCGCGGGGCGG 9340
Qy 9001 GACATTTATCAAGCGGTGTGCGCGGCCGACCCCGCTTATGCTCTTGGCTACTCTTA 9060
Db 9341 GACATTTATCAAGCGGTGTGCGCGGCCGACCCCGCTTATGCTCTTGGCTACTCTTA 9400
Qy 9061 CTTTTTGTAGGGGTAGGCTTTTCTTCACTCCCGCTCGG 9099
Db 9401 CTTTTTGTAGGGGTAGGCTTTTCTTCACTCCCGCTCGG 9439

RESULT 9
AB047641
LOCUS AB047641 9659 bp RNA linear VRL 04-AUG-2001
DEFINITION Hepatitis C virus gene for polyprotein, complete cds, clone:JCH-2.
ACCESSION AB047641
VERSION AB047641.1 GI:13122265
KEYWORDS
SOURCE
ORGANISM
Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.

REFERENCE
1 Kato, T., Furusaka, A., Miyamoto, M., Date, T., Yasui, K., Hiramoto, J.,
Nagayama, K., Tanaka, T. and Makita, T.
Sequence analysis of hepatitis C virus isolated from a fulminant
hepatitis patient
J. Med. Virol. 64 (3), 334-339 (2001)
JOURNAL
MEDLINE
21316767

PUBMED
REFERENCE
AUTHORS
JOURNAL

11424123
2 (bases 1 to 9659)
Kato, T., Wakita, T. and Furusaka, A.
TITLE
Submitted (23-AUG-2000) Takano Kato, The Tokyo Metropolitan
Institute for Neuroscience, Department of Microbiology; Mutsaersdai
2-6, Fuchu, Tokyo 183-8526, Japan (E-mail: takatoc@min.ac.jp,
Tel:81-423-25-3881(ex.4605), Fax:81-423-21-8678)
location/Qualifiers

FEATURES
source

1..9659
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/db_xref="taxon:11103"
/clone="JCH-2"
/note="Isolated from chronic hepatitis patient/genotype
2a"
341..9442
/codon_start=1
/product="polyprotein"
/protein_id="BAB32874.1"
/db_xref="GI:13122266"
/translation="MSTTPKPKTKNTNRRPODVKPEGGQIVGVVLLPRGRPL
GVRATKTSERSOPRRGRORI KDRRSRGKSGMKRGVPMPLVGRNIGAGMGLSPG
SRPMGPNDRPRRSRNKVIIDITLCGRADMGYIPVGAFLGVNAAALAGRVLED
GVNFRATGMLPGCSFSIFLLALISCTTPVSGAQRNTSSITVANDCANSTIMQLDA
AVLHVPGCVPEKVDTSRCWIPVSPNVAIPGALTRGLTHIDMVMSATLCSALY
VGLDCGVMLAAQMFIVSPORHMFVQBCNSIYPGHVSGHMAAMMMNNSPTATWL
AYVMRPEVILIDISGAHMGVFGLAIFSGMGAKAKVVLILLAGVDAGTAVAGST
AHTTKIASFPSPGARNIOLINTNGSMHRTALNCBSLNTGFLAIPAFANFNS
GCPERLSACRIEDPRIGMTLOYBNNTNSEDMPKPCNTYPPQCIIVARSICGY
YCFTPSPVVGITDRLGVPTTWSNEDVFLNSTRPSGSSCTMMNSTTKTC
GAPCRIPADPNASADLCPTDCFRKNDAAVFRGSGPMITPCLVDYPRHMYHC
TVNNTYFKIRMYVGVGVEHRLAACNFTGDRCDLEDSDRLSLSTTEMLPCT
YSDLPALSTGLHLHONIVDOYVNGSLPALTKYIVRENVLLFLLLADRVACAW
MLILGOEALAEKLVLAHSAASACNGELFVPIPAVAADIKRAPIATVSLTGLM
SFCULLALPOOAYAYDTSAHGOIGATLILITITPTLTPAKTIVSFYFWLCTLLT
GEANQEMAPPLQVRGKDGIIIMAAIFPGVFDITKMLALGPVLYLSDALTRVP
YFVRHALLRICTVRLHLAGRVYQOMLALGRWTGYIYDHLSPKMDMANGRLDA
VAVEPIIPSPMEKVIWGAETAAAGDILHGLPVSAIRGRLVAGADYSGKMSLL
APITAYAGOTGLLGAIVSMGTGRDKEOAGEIOVLSTVQSFLGTSLAGLVTVYSG
AGNKTLAGSRPTOMVSAAGDLEVMSPTGSLPCTCGAADVILVLRNALTIV
LRDGRKALISPRPLSTLKSSSGSPVLCRPGHAGVIFRAVCSRGVAKSIDIPVET
LDIVTRSPEDNSTPPVAPOTYOVGLHAFVSGSKTKVAAQYKVLVAPBV
AATLPGAVLISKAGINPNIRGTAVTVTAALITVYTKFLVAGGSGGAVYDIIID
ECHSVADATIIIGITVLDQAEATAGRSVLTATPQSVTTPHNIIEVLAGQGEIP
FYGRALPLSYIKGGHLLFCHSKKCDLAAALRGMGILNAVAYRGDVSIIIPQGVY
VVVATDALMTGTDPSVIDCNVAVTVQVDFSLDPTTITOTVPODVAERSGRBR
TGRGLGTIVRYVSTGERBASGMFDSVLSSECDGAAMTELTPATTVTRIVRTPBL
PVCODHLEFMEAVFTGLTHIDHFLISOTKAGDNPAITAYCAVCAKAPRPSMIV
MMKCLTRLKPTLVGPTPLVRLGVTNVTLLTHPTKTIATCOMADLEWMTSTVAVG
GILAAVAVYCLATGCVSIIIGRLINQRAVVA PDKEVLEAEDMBECASRALITEBQ
RIAEMLKSKIOGLLOQASKOADIOPAVQAMPVYQWAGMNNFTSGIYLAGLST
LPGNPAVSMMAFSAALTPSTSTTLLNTLIGMLAIOPLPGAIGFVYSGVGA
VCSIGTIGVLDIILAGVAGISGALVAKINSGRKMYEDVSLIPGLISGLLVVY
ICATILRRHVGBGAVVMMRLIAPASRGHVAPTHVYEDVLSQRYTGLASITTT
SLRLHNMITEGCPICPSGSMLRDVMWVCTIILDFENMLTSLKLPMPRLPVSQ
KGYKGVMAGTIMTTRPCGANISGNVAGSMRTGPCTCNATVQGLFVINCYGGC
VPRAPNFKTAIWRVAABEVAETOHGYSYITGLTNTLNVPCQLSPSEFSSVDYQ
OIHRAPTPPKPPFDEVSFCVGLNSFVVGSQLPCDPRBEGDMLDPBTHIAEFA
ABRLARGSPBSBASASASOISAPSLATCTTHGKTHDMDWALFMGQDVLTIESS
KYVVDLSLDPMYEEBGLDPSIPEETMLPKKFPFALPMPARPDYNPPLVSRKPY
QPTVAGCALPPPKTTPPPRRRTVGLSSBTIGDALQOLAIAFGQPPQGGSEVAP
TGADADAGSRTPPDESALSTGSISSMPPLEGREGDLDPRVBERPPQGGSEVAP
GSDGSMSTGSEDDSVCCSMYSMTALITPCSPBEKRIPIPLSLSHYNNKY
CTTSKASILRAKVTDFDMOYLDAYDSVLKDIKAASKVAKLITLLEBAOLTPHS
ARSKYGPAAKVRSLSGRAVNHISVMDLLEDSOTPTTITAKNEVFCYDPKGGK
KAARLIVPDLSVRCEKMALIDVQKIPQAVMGASVGFQSPORVEFLAKAAEKK
DPMGFSYDRCDDTVTERDITRTEBIVRAOSLPEEATIAHSLTERIVYGGPMFNK
GEOBCYRCSRASGVLTSMGNTITCYAALAAAKRAAGVATVACGDDILVISESG
TGSDBNLRFAETEAMTRYAPAGDPRPEYDELTGSEAVNSVALGCGGRRVYLR
DPTPIAARAEVTRHSPANMIGNIOYAPTIMARVHMLTPHSIIAODTLAONS
FEMGAVVSVSRFLDIPATIERLHGLDASLHTYSPHELTRVASSLRTKAGPLAAMS
RAVAVRASILSRGRAVCGRLFMNAVTKLKLPLPEARLLDSFVAGAGGDI

ORIGIN		YHVSRRARBRLLLLLSLLLSVGVGFLLPAR"									
Query Match		88.2%; Score 8025.4; DB 14; Length 9659;									
Best Local Similarity		92.6%; Pred. No. 0;									
Matches 8428; Conservative		0; Mismatches 671; Indels 0; Gaps 0;									
Qy	1	ATGAGCACAAAATCTTAACTTAACCTCAAGAAAAAACCAGAAAGAACCAACCGTCGCCACAA	60								
Db	341	ATGAGCACAAACTCTTAACTTAACCTCAAGAAAAAACCAGAAAGAACCAACCGTCGCCACAA	400								
Qy	61	GAGGTTAAGTTTCCGGGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	120								
Db	401	GAGGTTAAGTTTCCGGGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	460								
Qy	121	GGCCCAAGGTGGGTGGCGCGCGCAAGAAAGACTTTCGAGCGGCTCCAGCCACGTGGA	180								
Db	461	GGCCCAAGGTGGGTGGCGCGCGCAAGAAAGACTTTCGAGCGGCTCCAGCCACGTGGA	520								
Qy	181	AGGCGCCAGCCCATCCCTTAAAGATCGGCGCTCCACTGGCAAAATCTTGGGAAAAACAGGA	240								
Db	521	AGGCGCCAGCCCATCCCTTAAAGATCGGCGCTCCACTGGCAAGTCTTGGGAAAAACAGGA	580								
Qy	241	TACCCCTGGCCCTATACGGGAATGAGGCACTCGCTGGGCAAGATGGCTCTCTCTCC	300								
Db	581	TACCCCTGGCCCTATACGGGAATGAGGCACTCGCTGGGCAAGATGGCTCTCTCTCC	640								
Qy	301	CGAGTTCCCGTCCCTCTTGGGGGCCCAATGACCCCGGCAATAGTGGCGCAACGTGGGT	360								
Db	641	CGAGTTCCCGTCCCTCTTGGGGGCCCAATGACCCCGGCAAGTGGCGCAACGTGGGT	700								
Qy	361	AAAGTCATGATACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAAC	420								
Db	701	AAAGTCATGATACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAAC	760								
Qy	421	GGGCGCCCGTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	480								
Db	761	GGGCGCCCGTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	820								
Qy	481	GGGGTAAATTTTCAACAGGGAACCTTAACCGGTTGCTCTTTCTATTCTTCTTGCGGC	540								
Db	821	GGGGTAAATTTTCAACAGGGAACCTTAACCGGTTGCTCTTTCTATTCTTCTTGCGGC	880								
Qy	541	CTGCTGTCTGATCAACACCCCGGTCTCGCTGCGGAAAGTGAAGAACATCACTACCGGC	600								
Db	881	CTGCTGTCTGATCAACACCCCGGTCTCGCTGCGGAAAGTGAAGAACATCACTACCGGC	940								
Qy	601	TACATGTAATGAAGATGCAACCAATGACAGATTAACCTGGAGCTCCAGGCTGCTGTC	660								
Db	941	TACATGTAATGAAGATGCAACCAATGACAGATTAACCTGGAGCTCCAGGCTGCTGTC	1000								
Qy	661	CTCCACGTCCTCCGGGTCGCTCCCGTGGAGAAAGTGGGGAATGATCTGAGTGTGGATA	720								
Db	1001	CTCCACGTCCTCCGGGTCGCTCCCGTGGAGAAAGTGGGGAATGATCTGAGTGTGGATA	1060								
Qy	721	CCGGTCTCAACGAATGTCGCGTGAAGCGGCGCGCGCTTCAACGAGGCTTGGCGAGC	780								
Db	1061	CCGGTCTCAACGAATGTCGCGTGAAGCGGCGCGCGCTTCAACGAGGCTTGGCGAGC	1120								
Qy	781	CACATGCAATGTTGTGATGTCGCGCAAGCTTGTCTGCGCTTCAACGAGGCTTGGCGAGC	840								
Db	1121	CACATGCAATGTTGTGATGTCGCGCAAGCTTGTCTGCGCTTCAACGAGGCTTGGCGAGC	1180								
Qy	841	TGCGTGGGGTGAATGCTCGAGCGCAAAATTTCAATTGTCTCGCGGAGCACTGATTT	900								
Db	1181	TGCGTGGGGTGAATGCTCGAGCGCAAAATTTCAATTGTCTCGCGGAGCACTGATTT	1240								
Qy	901	GTCGAAGATGCAATTTGCTCATCTCACTGATGATCACTGATGATCACTGATGATCACTGAT	960								
Db	1241	GTCGAAGATGCAATTTGCTCATCTCACTGATGATCACTGATGATCACTGATGATCACTGAT	1300								
Qy	961	GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1020								

Db 1301 GACATGATGATGAACTGTGTCGCCCAAGGCTACCATGATTCGAGGCTACGTGATGCGGCTC 1360
Qy 1021 CCCGAGGTCAATTATGACATCATTAAGCGGGCTCATTTGGGGCGTCAATGTTGGCTTGGCC 1080
Db 1361 CCCGAGGTCAATTATGACATCATTAAGCGGGCTCATCGGGGCGTCAATGTTGGCTTGGCC 1420
Qy 1081 TACTTCTCTATGACGAGGAGCGTGGGCGAAAGTCGTTGTCATCCCTCTCTGTTGGCCGCGGG 1140
Db 1421 TACTTCTCTATGACGAGGAGCGTGGGCGAAAGTCGTTGTCATCCCTCTCTGTTGGCGG 1480
Qy 1141 GTGAGCGCGGCAACCCATCTGTTGGGAGTTCGCGCGCAAGACCAACCGGCGCTTACCC 1200
Db 1481 GTGAGCGCGGCAACCTATGCTGTGGGAGTTCACCGCGCACCAACCAAGGATCGCC 1540
Qy 1201 AGCTTATTTGACATGGGCCCCAGGAGAAATCCAGCTCGTTAACACCAATGGAGCTGG 1280
Db 1541 AGCTTCTTCTCCCTGGCGCGCCGAGAAATCCAGCTCATTTAACACCAATGGAGTGG 1600
Qy 1261 CACATCAACCGCAACCGGCTGAACCTGCAATGACTCTTGCAACACCGGCTTTATCGGCT 1320
Db 1601 CACATCAACCGCACTGCGCTGAACTGCAATGAACTCCCTGAACACCGGCTTTCTTGCGGC 1660
Qy 1321 CTGTTCTACACCAACAGCTTCAACTCGTCAGAGATGTCGCAACGATGTCGCGCTGCGC 1380
Db 1661 CTGTTCTACGCAACCGCTTCAACTCGTCAGAGATGTCGCAACGCTGTCGCGCTGCGCT 1720
Qy 1381 AGATTCAGAGCCTTCCGGGTGGAGTGGGGCGCTTGCAATTTGAGATTAATGTACCAAT 1440
Db 1721 AATATCAGAGACTTCCGGATAGGATGGGGGCAATTTGCAATACGAAATTAATGTACCAAT 1780
Qy 1441 CCAGAGATATGAGAACCTATTTGCTGACATACCCACCAAGGAGTGGGCTGCTGCC 1500
Db 1781 TCAGAGATATGAGAACATTTGCTGATCTAACCCACCAAGAGATGTTGGCATATGTCGCC 1840
Qy 1501 GCGAAGACTGTGTGTGGCCCAAGTGTACTGTTTACCCCGCAAGCCGAGTGTATGGGACG 1560
Db 1841 GCGAGGTCTGTATGTGGCCCAAGTGTACTGTTTACCCCGCAAGCCGAGTGTATGGGACG 1900
Qy 1561 ACCGACAGGCTTGGAGGCGCCACTTACAGTGGGGGAGAAATGAGACAGATGTTCTTA 1620
Db 1901 ACCGACAGACTCGGAGTGTACTTACAGTGGGGGAGAAATGAGACAGATGTTCTTA 1960
Qy 1621 TTGAACAGCACTGACACACCGGCTGGGCTCATGTTGGGCTGACGTTGAGAACTCTTCT 1680
Db 1961 TTGAACAGCACCGGACACCGTGGGCTCTGTTGGGCTGACGTTGAGAACTCTTCACT 2020
Qy 1681 GGCCTACACCAAGACTTGGCGGCGCAACCCCTGCGTACTAGAGCTGACTTCAACGCCAGC 1740
Db 2021 GGTTCACCAAAACATGCGGAGGACCACTTGGCGGATTAAGCTGACTTCAACGCCAGC 2080
Qy 1741 ACCGACCTGTTGTGCCCCAGGACTGTTTAAAGAAAGCATCTGATACCCTTACTTCAAA 1800
Db 2081 GCGGACCTGTTGTGCCCCAGGACTGTTTAAAGAAAGCATCTGACGCGGCTTACTTCAAA 2140
Qy 1801 TGGGAGCTGTGGGCGCTGCTCAACGCGCAAGGTCCTGATGCACTCCCTACAGGCTCGG 1860
Db 2141 TGTGGCTGTGGGCGCTGCTCAACGCGCAAGTGTGATGCACTCCCTACAGGCTCGG 2200
Qy 1861 CATTACCCCTGCAACAGTTAACTATACCATCTTCAAAATTAAGATGATGTGGAGAGGTT 1920
Db 2201 CATTACCCCTGCAACAGTTAACTATACCATCTTCAAAATTAAGATGATGTGGAGAGGTT 2260
Qy 1921 GAGCACAGGCTCACAGGCTGCAATTTTCACTGTGGGAGTCTTGCAACTTGGAGAGC 1980
Db 2261 GAGCACAGGCTGTGGGCTGCAATTTTCACTGTGGGAGTCTTGCAACTTGGAGAGC 2320
Qy 1981 AGAGACAGAGTCAACTGTCTCTTGTGTCATCCACACGAAATGGGCAATTTTACCT 2040
Db 2321 AGAGACAGAGTCAACTGTCTCTTGTGTCATCCACACGAAATGGGCAATTTTACCT 2380
Qy 2041 TGCTCTTACTGCACTGCGCGCTGTGCACTGCTTCTTCACTTCAACCAAAACATC 2100
Db 2381 TGCTCTTACTGCACTGCGCGCTGTGCACTGCTTCTTCACTTCAACCAAAACATC 2440

Qy 2101 GTGAGCTACAAATTCATGATGAGCTTATCACTGGCCCTCACAAAAATACATGTCGCCAGT 2160
Db 2441 GTGAGATGTAACAATAAGATGATGAGCTGTCACTGCTCTTCAAAATACATGTCGATGG 2500
Qy 2161 GAGTGGGTAAATCTTATTCCTGCTTATGCGGAGCCGAGGAGTTGCGGCTGCTTATGG 2220
Db 2501 GAGTGGGTAAATCTTATTCCTGCTTATGCGGAGCCGAGGAGTTGCGGCTGCTTATGG 2560
Qy 2221 ATGCTCATCTTGTGGGCGACGCGGAGAGACAGCACTAGAGAAAGCTGTATCTTGCAGCT 2280
Db 2561 ATGCTCATCTTGTGGGCGACGCGGAGAGAGCACTAGAGAAAGCTGTATCTTGCAGCT 2620
Qy 2281 GCGAGCGGCTAGCTGAGTCAATGAGTCTTATTTTGCATCTCTTTTGTGGGCTGG 2340
Db 2621 GCGAGCGGCTAGCTGAGTCAATGAGTCTTATTTTGCATCTCTTTGTGGGCTGG 2680
Qy 2341 TACATCAAGGAGTGGGAGTACCCCTTATGCTACCTATCCCTCACTGAGCTGTGCTT 2400
Db 2681 TACATCAAGGAGTGGGAGTACCCCTTATGCTACCTATCCCTCACTGAGTGTGCTT 2740
Qy 2401 AGCTTACTGCTCTAGCATTCGCCCAACAGGCTTATGCTATGACGATCTGTGATGCG 2460
Db 2741 TGCTTACTGCTCTGAGCATTCGCCCAACAGGCTTATGCTATGATGATGATGCTGCGCAGGA 2800
Qy 2461 CAGATAGAGCGGCTGCTGCTGATATGATGATGATGATGATGATGATGATGATGATGAT 2520
Db 2801 CAGATAGAGCGGCTGCTGCTGATATGATGATGATGATGATGATGATGATGATGATGAT 2860
Qy 2521 ACCCTTCTCAACCGGTTTTTGTGGTGTGTGCTATCTTGTGACCTTGGGAGAACTATG 2580
Db 2861 ACCCTTCTCAACCGGTTTTTGTGGTGTGTGCTATCTTGTGACCTTGGGAGAACTATG 2920
Qy 2581 GTCCAGAGTGGGACCACTATGACAGTGGCGGCTGCGGCTGATGCGATCATATGAGCC 2640
Db 2921 GTCCAGAGTGGGACCACTTATGAGTGGCGGCTGCGGCTGATGCGATCATATGAGCC 2980
Qy 2641 GTGCGCAATTTCAACCGAGTGTGTGTTGACATAACCAAGTGGCTTGGCGGCTT 2700
Db 2981 GCGGCATATTTGCGCGGCGTGTGTGTGACATAACCAAGTGGCTTGGCGGCTT 3040
Qy 2701 GGGCTGCTTACCTCTTAAAGAGTCTTGTGACGCGCTGCGCTTCTGTCAGGCTCAC 2760
Db 3041 GGGCTGCTTATCTCTTAAAGAGTCTTGTGACGCGCTGCGCTTCTGTCAGGCTCAC 3100
Qy 2761 GCTTCTAGAGATGACATGACAGGCAAGGCAATCTGCGGAGGAGCTACATG 2820
Db 3101 GCTTCTAGAGATGACATGATGATGATGATGATGATGATGATGATGATGATGATG 3160
Qy 2821 GCGCTACTAGCCCTTGGCAAGTGTGACCTGACCTTATGATGACCACTCACCCCTATG 2880
Db 3161 GCGCTACTAGCCCTTGGTATGATGATGATGATGATGATGATGATGATGATGATG 3220
Qy 2881 TCGGATTTGGGCTGCTAGTGTGCTGCGGAGCTTGGCGGCTGCGCTTATGATGATG 2940
Db 3221 TCGGATTTGGGCTGCAATGAGCTGCGGAGCTTGGCGGCTGCGCTTATGATGATG 3280
Qy 2941 AGTCCGATGGAAGAAAGTCAATGTTGTGGGAGCGGAGACAGTCTGTTGGGAGATT 3000
Db 3281 AGTCCGATGGAAGAAAGTCAATGTTGTGGGAGCGGAGAGCGGCTGATGAGGAGATA 3340
Qy 3001 TTACACGAGACTTCCCGTGTGCGCGGACTTGTGTGAGGAGTCTCTTGGCGGAGCTAT 3060
Db 3341 CTACACGAGACTTCCCGTGTGCGCGGACTTGTGTGAGGAGTCTCTTGGCGGAGCTAT 3400
Qy 3061 GGCCTATACCTCAAGGAGTGAAGTCTTGTGCGCCCATCACTGCTTACGCGCACAGACA 3120
Db 3401 GGCCTATACCTCAAGGAGTGAAGTCTTGTGCGCCCATCACTGCTTACGCGCACAGAG 3460
Qy 3121 GGTGGCTTTTGGGACCAATGATGATGATGATGATGATGATGATGATGATGATGATG 3180
Db 3461 GGTGGCTTTTGGGCGCTATATGATGATGATGATGATGATGATGATGATGATGATG 3520

3181 GGGGAAATTCAGGTCCTGTCTCCAGACGTCACCTAGTCCTCTCCGAAATTCATCTCGGG 3240
3521 GGGGAAATTCAGGTCCTGTCTACAGTCACCTAGTCCTCTCCGAAATTCATCTCGGG 3580
3241 GTTTTGTGACTGTCTACCTGAGAGCTGGCAACAAGACTTGGCCGGCTCAGGGGTCCG 3300
3581 GTCTTATGACTGTCTATCAATGGGGCTGGCAACAAGACTTACCTGCTCAACGGGTCCA 3640
3301 GTCAAGAGATGACTCTCAGTGTGAGAGGGGAGCTTATAGTGGGCCAGGCCCTGGG 3360
3641 GTACCGCAATGATCTCAATGTCGAGGGGAGACTTGTGAGATGGCCCTCCCTCGGG 3700
3361 ACTAAATCTTTGAGAGCCGTGACAGTGTGAGCCGTGACCTGTACTGTGACGCGGAAC 3420
3701 ACTAAGCTTTGAGAGCCGTGACAGTGTGAGCCGTGACCTGTACTGTGACGCGGAAC 3760
3421 GCTGATGTCAATCCCGGCTCGAAGACGCGGGGACAAACGGGGAGCGCTACTCTCCCGAGA 3480
3761 GCTGAGGTCAATCCCGGCTCGAAGACGCGGGGACAAACGGGGAGCACTACTCTCCCGAGA 3820
3481 CCTCTTTCACCTTGAAGGGGTCTCTCAGAGAGCCCGGTGTATGCCCCAGGGGCCAGCT 3540
3821 CCACTTTTCACCTTMAAAGGGTCTCTCGAGGGGGGCGGTGCTTTGGCCCCAGGGGCCAGCT 3880
3541 GTGGAAGTCTTCCGGGAGCTGTGTGCTCTCGAGGGCGTGGCTAAGTCCATAGATTTTCATC 3600
3881 GTGGAAGTCTTCCGGGAGCTGTGTGCTCTCGAGGGCGTGGCTAAGTCCATAGATTTTCATC 3940
3601 CCGGTGAGACACTTCCAGATCTGCAACGCGGTCCCCACCTTTAGTACAAACACACACCA 3660
3941 CCGGTGAGACACTTCCAGATCTGCAACGCGGTCCCCACCTTTAGTACAAACACACACCA 4000
3661 CCGGTGAGACACTTCCAGATCTGCAACGCGGTCCCCACCTTTAGTACAAACACACACCA 3720
4001 CCGGTGAGACACTTCCAGATCTGCAACGCGGTCCCCACCTTTAGTACAAACACACACCA 4060
3721 AGCACAAGTCTCTGCGCATATGCTGCTCAGGGGTATTAAGTGTAGTGTAAATCCC 3780
4061 AGCACAAGTCTCTGCGCATATGCTGCTCAGGGGTATTAAGTGTAGTGTAAATCCC 4120
3781 TCGTGTGCTGCGCATATGCTGCTCAGGGGTATTAAGTGTAGTGTAAATCCC 3840
4121 TCGTGTGCTGCGCATATGCTGCTCAGGGGTATTAAGTGTAGTGTAAATCCC 4180
3841 AACTTATGAGACTGAGTCAAGACTGTGACAGACGCGGGGCCCATCAGTACTCCACATAT 3900
4181 AACTTATGAGACTGAGTCAAGACTGTGACAGACGCGGGGCCCATCAGTACTCCACATATC 4240
3901 GGCATAATCTCTGCGCATATGCTGCTCAGGGGTATTAAGTGTAGTGTAAATCCC 3960
4241 GGCATAATCTCTGCGCATATGCTGCTCAGGGGTATTAAGTGTAGTGTAAATCCC 4300
3961 GAATGCGATGCGTGTGACCTTACCAATCTCTTGGATCGGAACAGTCTTGTATCAAGCA 4020
4301 GAATGCGATGCGTGTGACCTTACCAATCTCTTGGATCGGAACAGTCTTGTATCAAGCA 4360
4021 GAGACAGCTGGGGTCAAGACTTACCTGTGCTTACAGACTACGCCCTTGGGTCAATGACA 4080
4361 GAGACAGCTGGGGTCAAGACTTACCTGTGCTTACAGACTACGCCCTTGGGTCAATGACA 4420
4081 ACCCCCCACCCCAATAGAGAGGTGGCCCTTGGAGGAGAGGGGAGATCCCTTCTAT 4140
4421 ACCCCCCACCCCAATAGAGAGGTGGCCCTTGGAGGAGAGGGGAGATCCCTTCTAT 4480
4141 GGGAGGGCGATTCCTCTGTCTTACATCAAGAGAGAGAGATCTGATCTTCTGCAATTC 4200
4481 GGGAGGGCGATTCCTCTGTCTTACATCAAGAGAGAGAGATCTGATCTTCTGCAATTC 4540
4201 AAGAAAAAGTGTGACAGCTCGCGGCGGCCCTTGGGGGTATGGGCTTGAATCAATGAGCA 4260
4541 AAGAAAAAGTGTGACAGCTCGCGGCGGCCCTTGGGGGTATGGGCTTGAATCAATGAGCA 4600
4261 TACTACAGAGGTTGACAGTCTCGTAATACCACTCAGGAGAGAGCGTATGATGCTCGCC 4320

4601 TACTACAGAGGTTGACAGTCTCGTAATACCACTCAGGAGAGAGTGTGTGCTCGCC 4660
4321 ACCGAGCCCTCATACAGGGGTATACGGGGCTTTGACTCCGTATCGACTGCAACGTA 4380
4661 ACCGAGCCCTCATACAGGGGTATACGGGGCTTTGACTCCGTATCGACTGCAACGTA 4720
4381 GCGGTCACTCAAGTTGTAGACTTCAAGTTTGAAGCCCACTTACATCAACACAGATTT 4440
4721 GCGGTCACTCAAGTTGTAGACTTCAAGTTTGAAGCCCACTTACATCAACACAGATTT 4780
4441 GTTCCCTCAAGAGCTGTCTCAGTACGACGCGCCGGGGTGTGACGGGTATGGGGAAGCTG 4500
4781 GTTCCCTCAAGAGAGCTGTCTCAGTACGACGCGCCGGGGTGTGACGGGTATGGGGAAGCTG 4840
4501 GGCATTTATAGGTATGTTTCACTGGTGAAGAGCTCAGAAATGTTTGAAGTGTAGT 4560
4841 GGCATTTATAGGTATGTTTCACTGGTGAAGAGCTCAGAAATGTTTGAAGTGTAGT 4900
4561 CTCTGTGAGTGTACGACGACGAGGGCGCATGGTATGAGCTCAACATCGAGACACCC 4620
4901 CTCTGTGAGTGTACGACGACGAGGGCGCATGGTATGAGCTCAACATCGAGACACCC 4960
4621 GTTCAAGCTCAGGGCTATTTCAACACGCGCGTGTGCTGTGTGCAACATCTTGAAG 4680
4961 GTTCAAGCTCAGGGCTATTTCAACACGCGCGTGTGCTGTGTGCAACATCTTGAAG 5020
4681 TTTTGGGAGGAGCTTTTCAACGCGCGTGTGCTGTGTGCAACATCTTGTGCAACAT 4740
5021 TTTTGGGAGGAGCTTTTCAACGCGCGTGTGCTGTGTGCAACATCTTGTGCAACAT 5080
4741 AAGCAATCGGGGAGAAATTTCCGATATCTTACAGCTTACAGGCTTACAGTGTGCTGAG 4800
5081 AAGCAATCGGGGAGAAATTTCCGATATCTTACAGGCTTACAGGCTTACAGTGTGCTGAG 5140
4801 GCCAAGCCCCCCCCCTGCTGCTGAGACGTCATGTGGAATGTTTGAATCTGACTCAAGCCC 4860
5141 GCCAAGCCCCCCCCCTGCTGCTGAGACGTCATGTGGAATGTTTGAATCTGACTCAAGCCC 5200
4861 AACCTGTGAGGGGCCCAACACTCTGCTGACCGCTTGGGCTGTGTTACCAAGAGTCAAC 4920
5201 AACCTGTGAGGGGCCCAACACTCTGCTGACCGCTTGGGCTGTGTTACCAAGAGTCAAC 5260
4921 CTCACACATCCCGTGTGAGAAATATCATGCGCACCTGTGATCAAGCCGACTTGAAGTCA 4980
5261 CTCACACATCCCGTGTGAGAAATATCATGCGCACCTGTGATCAAGCCGACTTGAAGTCA 5320
4981 ACCAGCAATGGGTCTTGGCAGGGGAGTCTTGGCGGGCGCTGCGCGGTATTTGCTGGGG 5040
5321 ACCAGCAATGGGTCTTGGCAGGGGAGTCTTGGCGGGCGCTGCGCGGTATTTGCTGGGG 5380
5041 ACCGGGTGTTTGCATCATTCGCGCGCTTGCACATTAACAGAGAGCGTGTGAGCGCG 5100
5381 ACCGGGTGTTTGCATCATTCGCGCGCTTGCACATTAACAGAGAGCGTGTGAGCGCG 5440
5101 GACCAAGAGTCTCTATAGAGCTTTTGAATGAGATGAGAGAAATGTGCTCTAGAGCGGT 5160
5441 GACCAAGAGTCTCTATAGAGCTTTTGAATGAGATGAGAGAAATGTGCTCTAGAGCGGT 5500
5161 CTCATTGAAGAGGGGAGCGGATAGCGGAGATGCTGAATCCAAGATCCAAGCTTATTTG 5220
5501 CTCATTGAAGAGGGGAGCGGATAGCGGAGATGCTGAATCCAAGATCCAAGCTTATTTG 5560
5221 CAGCAAGCTTCAAAACAGCTCAAGCATTAACAACCACTGTGAGGCTTCAATGGCCCAAG 5280
5561 CAGCAAGCTTCAAAACAGCTCAAGCATTAACAACCACTGTGAGGCTTCAATGGCCCAAG 5620
5281 GTAGAACAAATTCGGGCCCAACACATGTGGAATCTTCAATTAAGCGGCTCAATCTGCA 5340
5621 GTAGAACAAATTCGGGCCCAACACATGTGGAATCTTCAATTAAGCGGCTCAATCTGCA 5680
5341 GGCATATCAACACTGCGAGGAAACCTGCAATGACTTCAATGATGCGCTTCAATGCGCC 5400

Db 5681 GAACTGTCAACACTGCGAGGAAACCCGCTGTGGCTTCATGATGATGCAATTCAGTCCGCC 5740
Qy 5401 CTACACGATCCGCTGTCAACAAGACCACTATCTTCTCAACATTTTGGGGGGCTGAGCTA 5460
Db 5741 CTACACGATCCGCTGTCAACAAGACCACTATCTTCTCAACATTTTGGGGGGCTGAGCTG 5800
Qy 5461 GCATCCCAAAATYGCAACCAACCCGCGGGGCACTGGCTTCTGTGTCAAGTGGCTAGTGGGA 5520
Db 5801 GCGTCCCAAAATCGGGCCACCCGAGGGGCCACTGGCTTCTGTGTCAAGTGGCTAGTGGG 5860
Qy 5521 GCTGCGGTAGGCACTATAGCTTATAGTAAAGTGTCTAGTGAATCTTGGCAGGGTATAGT 5580
Db 5861 GCTGCGGTAGGCACTATAGCTTATAGTAAAGTGTCTAGTGAATCTTGGCAGGGTATAGT 5920
Qy 5581 GCGGGCAATTCGGGGGCTCTGCTGCAATTCAGATGATGCTGGCGGAGAGCCCTCATG 5640
Db 5921 GCGGGCAATTCGGGGGCTCTGCTGCAATTCAGATGATGCTGGCGGAGAGCCCTCATG 5980
Qy 5641 GAGATGTCTCAACCTTGTGCTGCAATTCCTGCTCGGGTGCCTTGTAGTGGAGTC 5700
Db 5981 GAGGACGTCTGCAACCTGCTGCTGCAATTCCTGCTCGGGTGCCTTGTAGTGGAGTC 6040
Qy 5701 ATCTGCGCGGCATTTCTGCGCGCAACGCTGGGACCGGGGGAAGGCGCTGCAATGATG 5760
Db 6041 ATCTGCGCGGCATTTCTAGCGCGCAACGCTGGGACCGGGGGAAGGCGCAATGATG 6100
Qy 5761 AATAGATCTATGCTTGTGCTTCTCAAGAGAAATACGCTGCGCCCACTACCTAGTACG 5820
Db 6101 AACAGGCTCATGCTTGTGCTTCTCAAGAGAAACACGCTGCGCCCACTAGTACG 6160
Qy 5821 GAGTGGATGCTGCGAGCGCTGTGACCAACTAATTGAGCTCTTACCATTAACAGAGCTG 5880
Db 6161 GAGTGGATGCTGCGAGCGCTGTGACCAACTAATTGAGCTCTTACCATTAACAGAGCTG 6220
Qy 5881 CTCAAGAACTCCCAACTGATTAATGAGGACTGCGCCCATCCATGCGGCGCTGCTG 5940
Db 6221 CTTAAGAACTCCCAACTGATTAATGAGGACTGCGCCCATCCATGCGGCGCTGCTG 6280
Qy 5941 CTGCGGATGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6000
Db 6281 CTGCGGATGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6340
Qy 6001 TCCAAATTTATTCCAAGATGCGCGGCTCCCTTGTCTCTGCTCAAAAAGGGTACAG 6060
Db 6341 TCTAAATTTATTCCAAGATGCGCGGCTCCCTTGTCTCTGCTCAAAAAGGGTACAG 6400
Qy 6061 GCGGTGTGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6120
Db 6401 GCGGTGTGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6460
Qy 6121 GCGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6180
Db 6461 GCGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6520
Qy 6181 CAGGGGACCTTTCCTATCAATTTGTACACGAGGGGCGAGTGCCTGCGGAAACCCGCGCA 6240
Db 6521 CAGGGGACCTTTCCTATCAATTTGTACACGAGGGGCGAGTGCCTGCGGAAACCCGCGCA 6580
Qy 6241 AACTTTAAGGTGCGCATCTGAGAGGTGGCGGCTCAGAGTACGCGAGGTGACGCAAC 6300
Db 6581 AACTTTAAGGTGCGCATCTGAGAGGTGGCGGCTCAGAGTACGCGAGGTGACGCAAC 6640
Qy 6301 GGGGCAATCACTATCAATTAACAGGACTGCACTGATTAATTTGAAAGTCCCTGCAACTA 6360
Db 6641 GGGGCAATCACTATCAATTAACAGGACTGCACTGATTAATTTGAAAGTCCCTGCAACTA 6700
Qy 6361 CCTCTCCGAGTCTTCTTCTGAGTGAAGGAGTGAAGTCAATAGTCTTTCAGGAGTCC 6420
Db 6701 CCTCTCCGAGTCTTCTTCTGAGTGAAGGAGTGAAGTCAATAGTCTTTCAGGAGTCC 6760
Qy 6421 CCGAAGCGGTTTTTCCGAGTGAAGTCTGCTTGTGCTGCTTAAATTCATTTGTGCTC 6480
Db 6761 CCGAAGCGGTTTTTCCGAGTGAAGTCTGCTTGTGCTGCTTAAATTCATTTGTGCTC 6820

Qy 6481 GGGTCCAGGCTTCTCTGCAACCTGAAACCCGACACAGAGTATGATGCTCAATGCTAACA 6540
Db 6821 GGGTCTAGGCTTCTCTGCAACCTGAAACCCGACACAGAGTATGATGCTCAATGCTAACA 6880
Qy 6541 GATCATCTATATCAAGGAGAGTCAAGCGCGGCTTATGAGCGGGGGTCAACCCCA 6600
Db 6881 GATCATCTATATCAAGGAGAGTCAAGCGCGGCTTATGAGCGGGGGTCAACCCCA 6940
Qy 6601 TCCGAGCAAGCTCTTCCGAGAGCAGCTATCCGACCAATGCTGCAAGCACTGCAAC 6660
Db 6941 TCCGAGCAAGCTCTTCCGAGAGCAGCTATCCGACCAATGCTGCAAGCACTGCAAC 7000
Qy 6661 ACCAGGCAAAAGCTATGATGAGAGATGAGTGAATGCTAATCCGTTCAATGGGGGGAT 6720
Db 7001 ACCAGGCAAAAGCTATGATGAGAGATGAGTGAATGCTAATCCGTTCAATGGGGGGAT 7060
Qy 6721 GTGACTCGAATGAGTCTGGGTCMAAGTGTGCTGCACTCTGCAACCAATGATG 6780
Db 7061 GTGACTCGAATGAGTCTGGGTCMAAGTGTGCTGCACTCTGCAACCAATGATG 7120
Qy 6781 GAAGAAAGAGCGACCTTGAGCTTGATACATCAGATATCATGCTTCCCAAGAGAG 6840
Db 7121 GAAGAAAGAGCGACCTTGAGCTTGATACATCAGATATCATGCTTCCCAAGAGAG 7180
Qy 6841 TTTCCACAGCTTTACCGGCTGCGGCAACGCTGATTAACCAACCGCTTGTGGAATG 6900
Db 7181 TTTCCACAGCTTTACCGGCTGCGGCAACGCTGATTAACCAACCGCTTGTGGAATG 7240
Qy 6901 TGGAAGAGGCGAGTTTCAACCGGCACTGTTGCGGGCTGTGCTTCTCTCTCTAG 6960
Db 7241 TGGAAGAGGCGAGTTTCAACCGGCACTGTTGCGGGCTGTGCTTCTCTCTCTAG 7300
Qy 6961 AAAACCCGAGCGCTCCCCAAGAGGCGCGGCAAGTGGGCTTAATGAGACTCAT 7020
Db 7301 AAAACCCGAGCGCGCTCCCCAAGAGGCGCGGCAAGTGGGCTTAATGAGACTCAT 7360
Qy 7021 GAGATGCGCTTCAACAGCTGCGCATTAATGCTTGTGCGAGCGCCCAAGGGGAGT 7080
Db 7361 GAGATGCGCTTCAACAGCTGCGCATTAATGCTTGTGCGAGCGCCCAAGGGGAGT 7420
Qy 7081 TCAAGCTTTCACAGGAGGCGGCTGCGGATTCGCGGAGTCAAGCGCTCTGATGAG 7140
Db 7421 TCAAGCTTTCACAGGAGGCGGCTGCGGATTCGCGGAGTCAAGCGCTCTGATGAG 7480
Qy 7141 TTGACCTTTCGAGAGAGTTCATCTCTTCAATGCGCCCTCGAGGGGAGCTTGA 7200
Db 7481 TTGACCTTTCGAGAGAGTTCATCTCTTCAATGCGCCCTCGAGGGGAGCTTGA 7540
Qy 7201 GATCCAGACCTGAGAGCTGAGAGTGAAGCCCAACCCCGCCGAGGGGGGAGTGA 7260
Db 7541 GATCCAGACCTGAGAGCTGAGAGTGAAGCCCAACCCCGCCGAGGGGGGAGTGA 7600
Qy 7261 GCTCCCGGCTCGAGCTGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7320
Db 7601 GCTCCCGGCTCGAGCTGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7660
Qy 7321 TGTGCTCATGCTCATCTCTGAGACCGGGCTTAAATCTCTTGAATCCGGAAG 7380
Db 7661 TGTGCTCATGCTCATCTCTGAGACCGGGCTTAAATCTCTTGAATCCGGAAG 7720
Qy 7381 GAGAAGTTACGATTAACCTTGAAGCACTCCCTGTTGGATATCAACAAGGTATC 7440
Db 7721 GAGAAGTTACGATTAACCTTGAAGCACTCCCTGTTGGATATCAACAAGGTATC 7780
Qy 7441 TGTACCAACAAGAGCGCTCACTAAGGCTTAAAGGTTAATTTGATAGATGCA 7500
Db 7781 TGTACCAACAAGAGCGCTCACTAAGGCTTAAAGGTTAATTTGATAGATGCA 7840
Qy 7501 GTGCTGAGCTCTATCAAGCTCAAGTCTTAAAGCACTTAAGTGAAGGCTTCAAGCTC 7560
Db 7841 GTGCTGAGCTCTATCAAGCTCAAGTCTTAAAGCACTTAAGTGAAGGCTTCAAGCTC 7900

Oy	7561	ACCGAAGGCTCCCTCACATGGAAGAGGCTTGCCAGTTAAACCCGACCCCATTCGCAAGA	7562
Db	7901	ACTGCAAAAGCTCTCCTACCTTGGAGAGGCAATGCAGTTAAACCCCACTCTGTCAGAA	7960
Oy	7621	TCATAATATGGGTTTGGGCTTAAGAGAGTCCGCAAGCTTGTCCGGAGAGGGCGTTAACAC	7680
Db	7961	TCCAAGTACGGGTTTGGGGCTTAAGAGAGTCCGCAAGCTTGTCCGGAGAGGGCGTTAACAC	8020
Oy	7681	ATCAAGTCCGTTGGGAAGACCTCTCGAAGGACTCAGAAAACACAAATTCGCACAAACATT	7740
Db	8021	ATCAAGTCCGTTGGGAAGACCTCTCGAAGGACTCAGAAAACACAAATTCGCACAAACATT	8080
Oy	7741	ATGGCCAAAAATGAGTGTTCTGCCTGAGACCCCAACCAAGGGGGCAAGAAAGACAGCTCGC	7800
Db	8081	ATGGCCAAAAATGAGTGTTCTGCCTGAGACCCCAACCAAGGGGGTAAGAAAGACAGCTCGC	8140
Oy	7801	CTTATTCGTTTACCTCGACCTCGGGCTCAGGGTCTCCGAGAAAGATGGGCCCTTATGACATT	7860
Db	8141	CTTATTCGTTTACCTCGACCTCGGGCTCAGGGTCTCCGAGAAAGATGGGCCCTTATGAGATGC	8200
Oy	7861	ACACAAAAAATCTCCCTCAGCGGGTGATGAGGGGCTTCTTATGGAATTCAGATTTCCCGCGT	7920
Db	8201	ACACAAAAAATCTCCCTCAGCGGGTGATGAGGGGCTTCTTATGGAATTCAGATTTCCCGCGC	8260
Oy	7921	CAGCGGTAAGAGTTTCTCTTGAAGCATGGGCGGAAAAAAGAGACCTTATGGGTTTTTTCG	7980
Db	8261	CAGCGGTAAGAGTTTCTCTTGAAGCATGGGCGGAAAAAAGAGACCTTATGGGTTTTTTCG	8320
Oy	7981	TATGATACCCGATGCTTTGACTCGAACCGTCACTGAGAGAGCATCAGGACTGAGAGATGCC	8040
Db	8321	TATGATACCCGATGCTTTGACTCGAACCGTCACTGAGAGAGCATCAGGACTGAGAGATGCC	8380
Oy	8041	ATATATCGGGCTGCTCTCTTGGCCCGAGAGAGGCCACACTGCATATCACTGCTCTAATCTAG	8100
Db	8381	ATATATCGGGCTGCTCTCTTGGCCCGAGAGAGGCCACACTGCATATCACTGCTCTAATCTAG	8440
Oy	8101	AGACTTTACGTGGAGGGGCTTATGTTCAAACGAAAGGGCCAAACCTGGGGTTAAGAGCCT	8160
Db	8441	AGACTTTACGTGGAGGGGCTTATGTTCAAACGAAAGGGGCGAGTCTGGGGTTAAGAGCCT	8500
Oy	8161	TGCCGCGCACAGGGGGTCTCACCACTAGCATGGGGAAACACATCACTGCTTACGTGAAA	8220
Db	8501	TGCCGCGCACAGGGGGTCTCACCACTAGCATGGGGAAACACATCACTGCTTACGTGAAA	8560
Oy	8221	GCTTTAGCGGCTTGAANAAGCTGCAGAGGATTAATCGCGCCACATGCTGTTATGCGGCGAT	8280
Db	8561	GCTTTAGCGGCTTGAANAAGCTGCAGAGGATTAATCGCGCCACATGCTGTTATGCGGCGAT	8620
Oy	8281	GACTTGGTTGTCATCTCAGAAAGCCAGAGGGACCCGAGAGGACCGAGCGGAACCTBAGAGCC	8340
Db	8621	GACTTGGTTGTCATCTCAGAAAGCCAGAGGGACCCGAGAGGATGAGCGGAACCTBAGAGCC	8680
Oy	8341	TTCAACGAGGCTATGACAGGTAATTCGCGCCCTCTGTGTGACCCGCCCGACCGAGAGTAT	8400
Db	8681	TTCAACGAGGCTATGACAGGTAATTCGCGCCCTCTGTGTGACCCGCCCGACCGAGAGTAT	8740
Oy	8401	GATCTGAGCTGATTAACATCTTGCTGCTCAATGTGTTGTGGGCTGGGGCCCAAGAGC	8460
Db	8741	GATCTGAGCTGATTAACATCTTGCTGCTCAATGTGTTGTGGGCTGGGGCCCAAGAGC	8800
Oy	8461	CGCCGCAATATCTACTCTGACAGAGACCCTTACACTCCAAATCGCCCGGGCTGCTGGGAA	8520
Db	8801	CGCCGCAATATCTACTCTGACAGAGACCCTTACACTCCAAATCGCCCGGGCTGCTGGGAA	8860
Oy	8521	ACAGTTAAGCACTCCCTGTCATTCATGAGCTGGGAAACATCATCAGTACGCCCGAC	8580
Db	8861	ACAGTTAAGCACTCCCTGTCATTCATGAGCTGGGAAACATCATCAGTACGCCCGAC	8920
Oy	8581	ATATGGGCTCGCATGCTCTGATGACACACTTTCTTCTCATCTCTCATGAGCCCAAGACACT	8640
Db	8921	ATATGGGCTCGCATGCTCTGATGACACACTTTCTTCTCATCTCTCATGAGCCCAAGACACT	8980
Oy	8641	CTGAGCCAGAACTTCAACTTTGAGATGTACGAGAGGATGACTCCGTAAGTCCCTTGGAC	8700

Db	8981	CTGGACCCAGAACTCCAATTTTGAGATGTATGGAGCGGTGTACTGCCGTAAGTCTTTGGAC	9040
Qy	8701	CTCCGACCTATTAATTGMAAGTTACATGGGCTTTGACGCTTTTCTCTGCACACATACACT	8760
Db	9041	CTCCGACCCATTAATTGMAAGTTACACAGGGCTGACGCTTCTCTCTGCACACATACACT	9100
Qy	8761	CCCCACGAACTGACACGCGGTGGCTTCAGCCCTCAGAAAACTTTGGGGGGCCACCCCTTCAGA	8820
Db	9101	CCCCACCAACTGACGCGGGTGGCTTCAGCCCTCAGAAAACTTTGGGGGGCCACCTTCACAGA	9160
Qy	8821	GCGTGAAGAAGCCGGGAGCGTAGTCAGTACGAGGACGTCCTCATCTCCCGTGGGGGGAGAGCG	8880
Db	9161	GCGTGAAGAAGTCCGGGAGCGGTGACGTACGAGACGTCCTCATCTCCCGTGGAGGAGAGCG	9220
Qy	8881	GCCGTTTGGCGGTGATATCTCTTCAATTGGGCGGTGAGAAGCAAGCTCAAACTCATCTCA	8940
Db	9221	GCCGTTTGGCGGTGATATCTCTTCAATTGGGCGGTGAAAACAAAGCTCAAACTCATCTCA	9280
Qy	8941	TTGCGCGAAGCGCGCCTCTCGATTTATCCAGCTGATTCACCGTCCGCGCGCGGGGGGC	9000
Db	9281	TTGCGCGAAGGCGCGCCTCTCGATTTATCCAGCTGATTCACCGTCCGCGCGCGGGGGGC	9340
Qy	9001	GACATTTATACACAGCGTGTGCGCGTCCCGAACCCCGCTTATTGCTCTTTGGCTTACTCTTA	9060
Db	9341	GACATTTATACACAGCGTGTGCGCGTCCCGAACCCCGCTTATTACTCTTAACTCACTCTTA	9400
Qy	9061	CTTTTGTAGGGGTAGGCGCTTTTCCATCTCCCGGCTCCG	9099
Db	9401	CTTTCCGTAGGGGTAGGCGCTTTTCCATCTCCCGGCTCCG	9439

LOCUS	AF238481	9416 bp	RNA	linear	VRL 26-MAR-2000
DEFINITION	Hepatitis C virus 2a polyprotein gene, complete cds.				
ACCESSION	AF238481				
VERSION	AF238481.1	GI:7329200			
KEYWORDS					
SOURCE	Hepatitis C virus				
ORGANISM	Hepatitis C virus				
REFERENCE	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.				
AUTHORS	1 (bases 1 to 9416) Iwakura,J., Nagayama,K., Enomoto,N., Kurosaki,M., Watanabe,H. and Sato,C.				
TITLE	Full length cDNA sequence of HCV genotype 2a				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 9416)				
AUTHORS	Sato,C., Nagayama,K., Enomoto,N., Kurosaki,M., Watanabe,H. and Sato,C.				
TITLE	Direct Submission				
JOURNAL	Submitted (24-FEB-2000) Second Department of Internal Medicine, Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-Ku, Tokyo 113-8519, Japan				
FEATURES	location/Qualifiers				
source	1..9416				
	/organism="Hepatitis C virus"				
	/mol_type="genomic RNA"				
	/strain="MD2a-1"				
	/db_xref="taxon:11103"				
	/note="genotype: 2a"				
	307..9408				
	/codon_start=1				
	/product="polyprotein"				
	/protein_id="AA59940.1"				
	/db_xref="GI:7329201"				
	/translation="MSTNPKPQRTKANTRRPQNVKPPGGQIVGGVLLPRGRPL GVARTRTSRSQGRGRROPIDPDRSTGSGMCKPGPMPYLNREGIALAGWLLSPG SRPBGWGNDRHRSGRNVKYLIDLTGCFALMDVYIPVGAPLGVNARALAGHVRLED GVNFAFNTNLPQCSFSTFLMLLSCITIPVAQVQVKNYSNMYNTNCSNDSITMOWD AVLHVGCVPCVEKENTSRCKIVPSRVAARQAGALTONLRTHIDVWVSATICSALY VGDLGGVWLAQMFIIVSPRHWPFVECNOSIIPGIIIGHRMAMDMNMSPTTMTL				

AYAMRIPEVILIDIMSGAHNMVGLAYFSMOGAMAKVVILLITAGVDAETOTVGAT
GSSVLGFTSLFTRGPOOQIOLINTNGSMHINRFLNCONISIPKOGIIPABVYNNPSS
GROPMACRTIEAPRIGMGTLOYEDNNTNDEMRYCMTYPRKOGIIPABVYNNPSS
YCTPSPVYVVTGTDKLVPTTYGEMNTDVFLLNSTRPPRSGMFGCTMMNSTPTKTC
GAPRCRTRADFNASTDLCTPTDCFRKHPEATYIKCSGPMVTRCLVDYPRILMHPFC
TINYVFKRMVGVGVEHRLTAACNFRGRCDLBDREDOCPPLHSTTEMLIPCS
YSBLPLSTGLHLHONIVDVOYMGSLPALTXYVMEVWVLLIYAVACACVM
MLILGOAEALAEKLVILHAASACNGFIFYVIFVAAVYIGRAVPLAAYSLTGIM
PCLIVLALPOAVADASVHGXGAVLILITLITLTPITLTPYKTLRSCIMMLCYLLTL
GEMVNBWVPMQARGSDIIMANTIFCGGVFDTITKLLAVLGRVILRGLRNLTP
YFVRHALLMCTVHRHLAGRYVQMALALGRMTGYLYDHLTPMSDVAANLRLDA
VAVEPIIFSPMEKVIWGAETACGDILHGLPVSARLHGLIAGDGTSGMML
APTAAOQTRGLGAIIVSMTRGDKTEBGEOILSTYOSLIGLSAGVLTAVHG
AGNKTLAGSRGPVTOMYSSAEDLVGMPSPGKSLPCGCVADIVLYATRVADVI
PARGDRCGLLSPRLSTLKSGSGGVLCPRGHAGIIPAAVNSRVAASIDIPET
LDIVTRSPITSDNSTPPAVPOTYQVGLHAPTSKSTKFAVAAQKIVLANSV
AATLFGAALSKAHGINPNI RTGVRTVTGEPITYSYTGFLADGACAGAYOIIICD
ECHAADATTLIGIYVLDQAEVTRVLTATPPGSVTPPHNIEVLAEGEBIP
FYKAIPLSYIKGRHLIFCHSKKCDLALALRGMALNAAYRGLDVSIIPTQGV
VVVATDALMTGYGDFDTVDNVAIVOVVPSLDPITLITTOVPODASRORGR
TGRGLIYRYVSTGERASGMPDSVVLGECYDAGAAWELTPESTTVRLAYRNTGL
PVCQDLHEWAVFTGLTHIDAFISQTKQSGENFAYLVAYQIVCARAPAPSMV
MKCLIRLKEPVLGPPPLYRLGSVTNEVTLTHPKYIATCQADLEWTSYTWLAG
GLYAAVAAVCLATGVSIIGRLHINQAVAVAPKEVLYEAFDEMECASRALIEBQ
RLAEMLSKIQGLLOOASKOADIOPAVONAPKVEOPAKHMMNFISGLOYLAGST
LGNPNAVAMAFSAALTSPLSTTILNIIIGWLASOIAAPRAGTGFVSGLVGAA
VSGIGKVLVDMLAGYAGISGALVAFKIMSEKRSMDIVNLLRGLISRGLVYGV
ICAILIRHVGREBGAQVOMNRLIAPSRKNHAPHYTHYESASORVLOLSLIT
SLRLHNNMTEDCPIPCAGSMILRDVMDNLTLPKMLISKLPKPMKQPIICQ
KGRVWAGTGMTRPCPGANISGVNRLSMRTIGPKMTMOCQTPINCTEBCQ
VKPAPNFKTARIVAASEYAEVTOHGSYSITGLTDVNVKVCOLSPFPMVAVG
OIHFAPTPKPFRDVSFVSGLSVFSOLDEPDPDVLMSLTDPSHTAA
ARRLAGSRPSEASASOLASPLRATCTTHRADVDVMDNLMGCVTRTIEBS
KXVVIDSLDPMBERSDHEPSISEYLPESRPPLAPMARETNIPVLEPKRADY
QPPAVAGCALPPEKPTPTPEPRKPRVGNESITADLOQLAKRTGQPPSPDSGDS
TADADSGQTPPDELAHSBETGISMPLEBEPDPIESQVLESPPQGEVAP
GSDSGMSTCEBDDSVCCSMYSWTKALITCEPDEBEKLPIPLNSLSLRHANYK
CTTAKSASLRAKVTDFMOVDHAYHVKDKIILAASRVLILTEBAOCTLPHS
ASKXGPRKAEVRSLSGRAVNHKSKWIKLBESQPIPTTIAKREVCOTDPKSGK
KAARLIVYDGLVRCVKMALVDVTKLPQAVMGASGYOYSAQVLEFLKMAWK
DEWGFYDTRCEPDSTYTERDIRTESIYLAASLPEBARIAHSLTRLVYVGMFSK
GQTCGRASGVLTTSMGNTITCYVKAALACAKAIAVAPMLVCGDILVYSEEG
TEEDERILRAFTEMHRSYAPPGDPRPEDELLITSGSNVSLMOQDRRYLYLR
DPTPIIARAMEVTHRSVNSWMLGNTIOVAPTIWVMTWTHRSVILMOQDRYLYLR
PEMGSVYSVPLDPAIIRLHGLDAFSIHTYTPHELRIVASALKGLAPLRTKS
RAVAVASLISHGRLAICGRYLFNMAVTKLTPLEARLIDISSWTFVAGGSDI
YHSVSRARBRLLIFSLILSVGVLFLPLPAR"

ORIGIN

Query Match 86.8%; Score 7895.4; DB 14; Length 9416;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 8346; Conservative 0; Mismatches 752; Indels 0; Gaps 0;

QY 1 ATGAGCAAAATCTCTAACTCAAGAAAAAACCAGAAACCAACCGTCCGCCACAA 60
DB 307 ATGAGCAAAATCTCTAACTCAAGAAAAAACCAGAAACCAACCGTCCGCCACAA 366
QY 61 GACGTTAAGTTTCGGGCGCGCGCGCGAGATGTTGGCGAGATTAATCTTTGCCGCGAG 120
DB 367 AACGTTAAGTTTCGGGCGCGCGCGCGAGATGTTGGCGAGATTAATCTTTGCCGCGAG 426
QY 121 GGGCCGAGTTGGGTGGCGCGGACAGAGAAAGATTTCGAGAGGGTCCCGACGACGTGA 180
DB 427 GGGCCGAGTTGGGTGGCGCGGACAGAGAAAGATTTCGAGAGGGTCCCGACGACGTGA 486
QY 181 AGCGCGACAGCCATCCCTAAAGATCGGCGCTCCACTGCGCAAAATCTCGGGAAAAACAGGA 240
DB 487 AGCGCGACAGCCATCCCTAAAGATCGGCGCTCCACTGCGCAAAATCTCGGGAAAAACAGGA 546
QY 241 TACCCCTGGCCCTATACGGGAATGAGGGAATCGGCTGGCGACAGATGAGCTCTGTCCCC 300
DB 547 TACCCCTGGCCCTATACGGGAATGAGGGAATCGGCTGGCGACAGATGAGCTCTGTCCCC 606

QY 301 CGAGTTCGCCGTCTCTTGGGGGCCCCAATGATACCCCGGAGTAGTGCGGCAACGTGGGT 360
DB 607 CGAGTTCGCCGTCTCTTGGGGGCCCCAATGATACCCCGGAGTAGTGCGGCAACGTGGGT 666
QY 361 AAGGTATGATACCCCTAAAGATCGGCGCTTGGCGACCTCAATGGGGATACATCCGTCTGTG 420
DB 667 AAGGTATGATACCCCTAAAGATCGGCGCTTGGCGACCTCAATGGGGATACATCCGTCTGTG 726
QY 421 GGGCCCCCGTGGGGGGCGGCGGACAGAGCTCTGGCGCATGGCGTGAAGATCTGGAGAGC 480
DB 727 GGGCCCCCGTGGGGGGCGGCGGACAGAGCTCTGGCGCATGGCGTGAAGATCTGGAGAGC 786
QY 481 GGGGTTAATTTTCAACAGGAACTTACCAGTGTGCTCTTCTATCTTCTTGGCTGGCC 540
DB 787 GGGGTTAATTTTCAACAGGAACTTACCAGTGTGCTCTTCTATCTTCTTGGCTGGCC 846
QY 541 CTGCTGTCTGATCAACACCCCGGTCTCGCTGCGAAGTGAAGAACATCAGTACCGGC 600
DB 847 CTGCTGTCTGATCAACACCCCGGTCTCGCTGCGAAGTGAAGAACATCAGTACCGGC 906
QY 601 TACATGATGATCAACAGCTGACCAATGACAGCATTACTGGCAGCTCCAGGCTGCTGT 660
DB 907 TACATGATGATCAACAGCTGATGTTTCAACAGCAGCATTACTGGCAGGTTCAAGACGCGCT 966
QY 661 CTCACAGTCCCGGGGTGCGTCCCGTGGAGAAAGTGGGGAATGCATCTCAGTCTGGATA 720
DB 967 CTCACAGTCCCGGGGTGCGTCCCGTGGAGAAAGTGGGAATGCATCTCAGTCTGGATA 1026
QY 721 CCGGTCTCACGGAATGTGGCGGTGACAGCGGCGCGGCGCTCAACGAGGGCTTGGCAGC 780
DB 1027 CCGGTCTCACGGAATGTGGCGGTGACAGCGGCGCGGCGCTCAACGAGGGCTTGGCAGC 1086
QY 781 CATATGACATGATGTGATGTGCGGACAGCTCTGCTCTGCTTCACTGCGGGAGACTTC 840
DB 1087 CATATGACATGATGTGATGTGCGGACAGCTCTGCTCTGCTTCACTGCGGGAGACTTC 1146
QY 841 TGGGTGGGGGTGATGTGCGGAGCCAAATGTTTCAATGTCGCGCGACACCACTGGTT 900
DB 1147 TGGGTGGGGGTGATGTGCGGAGCCAAATGTTTCAATGTCGCGCGACACCACTGGTT 1206
QY 901 GTCCAAAGATGCAATTTGCTTCATCTACCTGTGACCATGACAGCAGCATGGCATGG 960
DB 1207 GTCCAAAGATGCAATTTGCTTCATCTACCTGTGACCATGACAGCAGCATGGCATGG 1266
QY 961 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 1267 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1326
QY 1021 CCGGAGTCAATTTAGACATCATTTAGCGGGGCTCATTTGGGGGCGTCATGTTGGCTGGCC 1080
DB 1327 CCGGAGTCAATTTAGACATCATTTAGCGGGGCTCATTTGGGGGCGTCATGTTGGCTGGCC 1386
QY 1081 TACTTCTATGACAGGAGGATGAGGAGGAGGATGATGATGATGATGATGATGATGATGAT 1140
DB 1387 TACTTCTATGACAGGAGGATGAGGAGGAGGATGATGATGATGATGATGATGATGATGAT 1446
QY 1141 GTGGAGCGCGACCCATATCTGTGGGGGTTCTGCGCGGACGACACCGGGCGCTTACC 1200
DB 1447 GTGGAGCGCGAAACCCAAACAGTGGGGGTGCAACCGGGGCGAGTCTTGGGGCTTACC 1506
QY 1201 AGCTTATTGACATGGGGCCCGAGGAGAAATACAGTCTGTTTAAACCAATGGCAGCTGG 1260
DB 1507 AGCTTATTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1566
QY 1261 CACATCAACCGACCGCGCTTGAATCTGCAATGATCTCTTGGACACCGGCTTATCGCGTCT 1320
DB 1567 CACATCAACCGACCGCGCTTGAATCTGCAATGATCTCTTGAACACCGGCTTATCGCGTCT 1626
QY 1321 CTGTTTCAACCAACGCTTCAATCTGTCAGAGTGTCCGGAAGCAAGTGTCCGCTGCGC 1380
DB 1627 CTGTTTCAACCAACGCTTCAATCTGTCAGAGTGTGTCCCAACGCAAGTGTCCGCTGCGC 1686
QY 1381 AGTATGAGGCGCTTCCGGGTGGAGTGGGGCGCTTTCGAATATGAGGATATATGTACCAAT 1440

Db	1687	ACCATCGAGGCGCTTTCGATATAGATGGGGCACTTGGCAATCGAGATTAATGTCCAAAC	1746
Oy	1441	CCAGAGATATAGAACCTTATTTGCTGGCACTACCCAGCAGTGTGGCGTGTCTCC	1500
Db	1747	CCAGAGATATAGAGCCATCTACGTGGGCACTACCCACCAAAACAATGTGGCATGTCCCC	1806
Oy	1501	GCGAAGATGTGTGTGGCCGAGTGAATCTGTTCAACCCAGCCAGTGTATGTGGGACG	1560
Db	1807	GCGAGTCAAGTGTGGCCCGAGTACTGTTTCACTCCAGCCCAAGTGTATGTGGGACG	1866
Oy	1561	ACCGACAGGCTTGGAGCGGCCACTTACAGTGGGGGAGATGAGACAGATGTCTTCTA	1620
Db	1867	ACAGATAACTTGGAGTGGCCCACTTACAGTGGGGCGAATGAGACAGATGTCTTCTA	1926
Oy	1621	TTGAACAGACTCCAGCCACCGCTGGGGTCAATGTTCGGCTGCACTGATGAACTCTTCT	1680
Db	1927	TTGAACAGACCCGACACCGCGGGGGTCAATGTTCGGCTGCACTGATGAACTCTCACT	1986
Oy	1681	GGCTTACCAAGACTTGGCGGCGACACACCTGCGTATAGAGCTGACTTTCACGCGACG	1740
Db	1987	GGCTTACCAAGACTTGTGGGCGACACCTTCCGCACTAAGCTGACTTTCACGCGACG	2046
Oy	1741	ACGAGCTGTGTGTGCCCAAGGACTGTTTATGAAAGCATCTGTATACCACTTACTTCAAA	1800
Db	2047	ACGAGCTGTGTGTGCCCAAGGACTGTTTATGAAAGCATCTGTAAAGCACTTATCATCAAA	2106
Oy	1801	TGGCGCTCTGGGGCCCTGGCTCACGCGCAAGTGGCTGATCGACTACCCCTACAGGCTCTGG	1860
Db	2107	TGGGTTCTGGGCGCTTGGCTCACGCGCAAGTGGCTGTTGATCTAACCCCTACAGGCTCTGG	2166
Oy	1861	CATTACCCCTGACAGATTAACTATACATCTTCAAAATTAAGATGTATGTGGAGGGGTT	1920
Db	2167	CATTACCCCTGACAGATTAACTATACACCGTCTTCAAGGTGAAGATGTATGTGGAGGGGTT	2226
Oy	1921	GAGCACAAGCTTCACGGGCTGCATGCAATTTCACTGTTGGGATTCGTTGCACTTTGAGAGAC	1980
Db	2227	GAGCACAAGCTTCACGGGCGGTGCATTTCACTGTTGGGATTCGTTGCACTTTGAGAGAC	2286
Oy	1981	AGAGACAGAAAGTCAACTGTCTCTCTTGTGTGCACTCAACAGGAATGGGCATTATTACT	2040
Db	2287	AGGACAGAAAGTCAACAGACTCTTGTGTGCACTCAACAGGAATGGGCCATCTTGTGCC	2346
Oy	2041	TGCTCTTACTGGGACCTGCGCGGCTTGTGCACTGTCTTCTTCACCTTCACCAAAACATC	2100
Db	2347	TGCTCTTACTGAGAAATGCGCGCTTGTGCACTGTCTTCTTCACCTTCACCAAAACATC	2406
Oy	2101	GTGACGATACAAATTATGATATGGCTTATCACTGCGCTCACAAAATACATGTCGCAATGG	2160
Db	2407	GTGACGATACAAATATATATATATGCTGTGCACTGCGCTTACAAAATATATGTGTCGGTGG	2466
Oy	2161	GAGTGGGATTAATCTTATTTCTGTGCTCTTAAAGGACGCGCAGGGTTTGGCGCTGCTTATGG	2220
Db	2467	GAGTGGGATTAATCTTATTTCTGTGCTCTTAAAGGACGCGCAGGGTCTGGCGCTGTGTATGG	2526
Oy	2221	ATGCTCATCTTGTATGGGCGAGGCGGACAGACACTAGAGAAAGTGTATATTTTGCACGCT	2280
Db	2527	ATGCTCATTTTGTGTGGGCGAGGCGGACAGACAGGCTAGAGAAAGTGTGTCTTGCATGCT	2586
Oy	2281	GCGAGCGGAGCTAGCTGCAATGTGCTCTTATTTTGTATCTTTTGTGAGCTGTGG	2340
Db	2587	GCGAGCGGCGGCTAGCTGCAATGTGCTCTTATTTTGTATCTTTTGTGAGCTGTGG	2646
Oy	2341	TACATCAAGGCTGGGATGTCCCTTAAAGTACTATTCCTCACTGGCGCTGTGTCTT	2400
Db	2647	TACATCAAGGCTGGGCGGCTCCCTTGGCGGCTATATTCCTCACTGGCGCTATGTGGCCCTTC	2706
Oy	2401	AGCCTACTGCTCTTACGATATGGCGGCAACAGGCTTATGTATATGACGATGTGTGATGAC	2460
Db	2707	TGCTCTACTGTCTTACGATGTGCGGCAACAGGCTTATGTATATGATGATCTGTGTGACGGA	2766
Oy	2461	CAGATAGAGCGGCTCTGTGTGTATGATCACTCTTAACTCTCAACCCCGGGTATTAAG	2520

Db	2767	CAGATNGCGGGTTCGTGTGATACATCTAATTAACCTCTTTACACTCACCCCGGTTATAG	2826
Qy	2521	ACCCTTCTACGCCGGTTTTGTGGTGTGTCTATCTTCTGACCTCGGGGAAAGTATG	2580
Db	2827	ACTCTTCTCAGCCCGTGCCTGTGTGTGTGTCTGTGTATCTCTGACCTCGGGGAAAGCATG	2886
Qy	2581	GTCAGAGAGTGGGACCAACCTATGACAGGTGGCGGGTGGCGGTATGAGCATATATGAGCC	2640
Db	2887	GTCAGAGAGTGGGATCCACCCATGACAGGACACGGGTGGCGGTATGAGCATATATGAGCC	2946
Qy	2641	GTCGCATATTTCTACCCAGGTGTGTGTGTGACATAAACAAATGGCTTTGGCGGTCTT	2700
Db	2947	GCCACCAATATTTCTCCCGGGGTGTGTGTGTGACATTAACCAAGTGGCTTCGGCCGTGCTT	3006
Qy	2701	GGGCGCTGTTACCTCCCTAAAGAGTCTTTGACGGCGCGTCCGTACTTTCGTCAAGGCTCAC	2760
Db	3007	GGCGCTGTTACCTCTTACAGAGTGTCTTTAGCGCGCGTCCCATCTTCTGTAGAGCTCAC	3066
Qy	2761	GCTCTACTGAGAGTGTGACCCATGGCAAGGCATCTCGCGGGGGCAGGTACGTCCAGATG	2820
Db	3067	GCTCTGCTGAGGATGTGACTGTGTGTGTGAGCACCTCGCGGGGGGTATGATGATGCAGATG	3126
Qy	2821	GCGCTACTAGCCCTTTGGACAGGTGGACCTTGACCTTAATCTATGACACCTCACCCCTATG	2880
Db	3127	GCGCTATTAGCCCTTTGGAGGTGACCGGCACTTAATCTATGACACCTCACCCCTATG	3186
Qy	2881	TCGGATTGGGCTGTGAGTGCGCTCGGGACCTCGCGGTGCGCGTGTGAGCCTATCATCTTC	2940
Db	3187	TCGGATTGGGCTGTGACCGGCTCGAGACCTTGGCGGTGCGCGTGTGAGCCTATCATCTTC	3246
Qy	2941	AGTCCGATGAGAGAAAGAAATGATTGTCTGTGGGAGCGGAGACAGCTGCTTGTGGGACATT	3000
Db	3247	AGTCCGATGAGAGAAAGAGTCAATGCTGTGGGAGCGGAGACAGCTGATGTGGGAGACATC	3306
Qy	3001	TTACACGGACCTTCCCGTGTCCGCCCACTTGTGTGGGAGGTCCTCTTGGCCGACGTGAT	3060
Db	3307	TTACACGGACCTTCCCGTATCCGCCCGACTCGGTGGGAGATTTCTCTTGTGGCCAGCTGAT	3366
Qy	3061	GGCTATACCTCCAAAGGGGTGGAGTCTTCTGCCCCCATCACTGTCTTACGCCACAGACA	3120
Db	3367	GGCTATACCTCTAAAGGGGTGGAGCTTCTGCCCCCCATCAACCGCTTACGCCACACAGACG	3426
Qy	3121	CGTGCCCTTTTGGGACCATAGTGTGTAGACATGACGGGGCGCGACAGACAAACAGGCT	3180
Db	3427	CGAGGTCTTTGGGGCGCATAGTGTGTACATGACATGACGGGGCGCGACAGACAAACAGGCT	3486
Qy	3181	GGGGAAATTCAGGTCTGTCCACAGTCACTCAGTCTCTTCTCGGAAACATCCATCTCGGGG	3240
Db	3487	GGGGAAATTCAAATCTGTGTCCACGGTCAACCAAGTCTCTTCTCGGAAACATCCATTCGGGG	3546
Qy	3241	GTTTTGTGAGCTGTCTACATGAGCTGGCAACAAACTCTGGCCGGCTCACGGGGTTCG	3300
Db	3547	GTCCTATGAGCTGTCTACACGGAGCTGGCAACAAACCTTAGCCGGCTCACGGGGCCCG	3606
Qy	3301	GTCACGAGATGTACTCCAGTGTGGAGGGGAACTTAGAGGTGGGCCAGGCCCCCTGGG	3360
Db	3607	GTCACGAGATGTACTCAAGTGTGGAGGGGAACTTGTGGGTGGGCCAGGCCCTCTGGG	3666
Qy	3361	ACTAATCTTTGAGCCGTGACAGTGTGGACCGGTGCACCTGTACTGTGTCAACGCGGAAAC	3420
Db	3667	ACCAATCTTTGAGCCGTGACAGTGTGGACCGGTGCACCTGTACTGTGTCAACGCGGAAAC	3726
Qy	3421	GCTGATGTCAATCCCGGCTCGAAGACCGGGGACAAACGGGGAGGGCTATCTTCCCGGAG	3480
Db	3727	GCTGATGTCAATCCCGGCTCGAAGACCGGGGACAAAGCGGGAGGGCTATCTTCCCGGAG	3786
Qy	3481	CCTCTTTCCACCTTGAAAGGGTTCCTCAGAGAGCCCGGTGTATGCCACGAGGGCCACGCT	3540
Db	3787	CCCCTTTGACCTTGAAAGGGGTTCCTCGGGGGGACCGGTGTCTTGGCCCAAGGGGACATGCT	3846
Qy	3541	GTCGAGTCTTCCGGGCACTGTGTGTCTCTCGGGCGTGTGCTTAAGTCCATAGATTTCAATC	3600
Db	3847	GTCGGGATCTTCCGGGCACTGTGTGTCTCTCGGGCGTGTGCTTAAGTCCATAGATTTCAATC	3906

3601 CCGGTGAGACATCTGACATGTCACGCGGTGCCCCACCTTTAGTGAACAAGACACCA 3660
3907 CCGGTGAGACATCTGACATGTCACGCGGTGCCCCACCTTTAGTGAACAAGACACCA 3966
3661 CTTGCTGTGCCCCAGACCTATCAGGTGCGGTACTTGTGATGCCCCGAGCTGGCAGTGAAG 3720
3967 CAGCTGTGCCCCAGACCTATCAGGTGCGGTACTTGTGATGCCCCGAGCTGGCAGTGAAG 4026
3721 AGCACAAAGTTCTGTGCGATATGCTGTGAGGGGTATTAAGTGTAGTGTCTTAATCCC 3780
4027 AGCACTAAGGTCTGTGCGATATGCTGTGAGGGGTATTAAGTGTAGTGTCTTAATCCC 4086
3781 TCAGTGCTGCGACCTGCGGGTTTGGGGCGTACTGTCTAAGGCAATGGCATCAATCCC 3840
4087 TCGGTGCTGCGACCTGCGGGTTTGGGGCGTACTGTCTAAGGCAATGGCATCAATCCC 4146
3841 AACATTAGACTGAGATCAAGACTGTGACGACCGGGGCGCCCATCACTACTCAATAT 3900
4147 AACATTAGACTGAGATCAAGACTGTGACGACCGGGGAGCCCATCACTACTCAATAT 4206
3901 GGCATAATTCCTCGCCGATGGGGGCTGTGCGGGCGGCGCTTACGATCATATATGAT 3960
4207 GGCATAATTCCTCGCCGATGGGGGCTGTGCGGGCGGCGCTTATGACATCATATATGAT 4266
3961 GAATGGCATCGCGTGAAGCTTACCAACATCCTTGGCAATGGGAAGTCTTGAACAGA 4020
4267 GAATGGCATCGCGTGAAGCTTACCAACATCCTTGGCAATGGGAAGTCTTGAACAGA 4326
4021 GAGACAGCTGGGGTCAAGATTAAGTGTGCTGCTACAGCTAGCCCCCTGGGTCACTGACA 4080
4327 GAGACAGCTGGGGTCAAGATTAAGTGTGCTGCTACAGCTAGCCCCCTGGGTCACTGACA 4386
4081 ACCCCCAACCCCAATAGAGAGAGTGGCCCTTGGGCAAGAGGGCGAGTCCCTTCTAT 4140
4387 ACTCCCATCTTAATAGAGAGAGTGGCCCTTGGGCAAGAGGGCGAGTCCCTTCTAT 4446
4141 GGGAGGGCGATTCCTCGTCTTACATGAAGAGAGAGAGATCTTGTGCAATTC 4200
4447 GGGAGGGCGATTCCTCGTCTTACATGAAGAGAGAGAGATCTTGTGCAATTC 4506
4201 AAGAAAAAGTGAAGAGCTGCGCGGGCCCTTGGGGTATGGCTTGAATCACTAGTGA 4260
4507 AAGAAAAAGTGAAGAGCTGCGCGGGCCCTTGGGGTATGGCTTGAATCACTAGTGA 4566
4261 TACTACAGAGGGTGAAGCTGCTCCGTAATACCACTCAGAGAGAGCTAGTGTGCTGCC 4320
4567 TACTACAGAGGGTGAAGCTGCTCCGTAATACCACTCAGAGAGAGCTAGTGTGCTGCC 4626
4321 ACCGACGCTCATGAAGAGGTATCTGGGAGCTTTGATCCTCGTATCGATCGAAGCTA 4380
4627 ACCGACGCTCATGAAGAGGTATCTGGGAGCTTTGATCCTCGTATCGATCGAAGCTA 4686
4381 GCGGTCACTCAAGTTGTAGACTTCAAGTTAGACCCCATTCACATTAACACACAGATT 4440
4687 GCGGTCACTCAAGTTGTAGACTTCAAGTTAGACCCCATTCACATTAACACACAGATT 4746
4441 GTCCCTCAAGACGCTGTCTCAAGTACGACGCGCGGGGTGCAAGGTTAGGGGAAAGCTG 4500
4747 GTCCCTCAAGACGCTGTCTCAAGTACGACGCGCGGGGTGCAAGGTTAGGGGAAAGCTG 4806
4501 GGGATTTATAGGTATGTTTCACTGAGTACGAGCTCAGGAAGTTTGAAGTGTAGT 4560
4807 GGGATTTATAGGTATGTTTCACTGAGTACGAGCTCAGGAAGTTTGAAGTGTAGT 4866
4561 CTCTGTGAGTGTCAAGACGAGGGGCGCATGTATAGCTCACAACCATGAGAACCA 4620
4867 CTCTGTGAGTGTCAAGACGAGGGGCGCATGTATAGCTCACAACCATGAGAACCA 4926
4621 GTCAAGCTCAGGGGCTATTTCAACAGCGCGGGTTTGGCTGTGTGCAAGACATCTTGA 4680
4927 GTCAAGCTCAGGGGCTATTTCAACAGCGCGGGTTTGGCTGTGTGCAAGACATCTTGA 4986

4681 TTTTGGAGGAGTTTTCACCGGCTCACAACAATAGATGCCCATCTCTTTCCCAACA 4740
4987 TTTTGGAGGAGTTTTCACCGGCTCACAACAATAGATGCCCATCTCTTTCCCAACA 5046
4741 AAGCAATCGGGGAAAAATTTGCAATCTTAACAGCTTACAGAGGTCAAGTGTGCTTAA 4800
5047 AAGCAATCGGGGAAAAATTTGCAATCTTAAGTGTGCTTACAGAGGTCAAGTGTGCTTAA 5106
4801 GCCAAAGCCCCCCCCCTGCTGTGAGAGTCAATGTGAAGTGTGACTTCACTCAAGCC 4860
5107 GCCAAAGCTCTCTCCCGTCTGGAAGTATGTGAAGTGTGATTTCACTCAAGCC 5166
4861 ACACTGTGGAGCCCAACCTCTCTGTACCGCTGTGGGCTCTGTATCAACAGAGTCA 4920
5167 ACCGTGTGGAGCCCTTACACCTCTCTGTACCGCTGTGGGCTCTGTATCAACAGAGTCA 5226
4921 CTCAACATCCCGTGAAGAAATATATGACCACTGACAGAGCCGACTTGAAGTCA 4980
5227 CTCAACATCCCGTGAAGAAATATATGACCACTGACAGAGCCGACTTGAAGTCA 5286
4981 ACCAGCAATGGGTCTTGGAGAGGGAGTCTTGGCGGCGCTGCGCGGTATGTGCTGGCG 5040
5287 ACCAGCAATGGGTCTTGGAGAGGGAGTCTTGGCGGCGCTGCGCGGTATGTGCTGGCG 5346
5041 ACCGGGTGTGTGATCATATCGGCGGCTTGCATTAACAGAGAGCGCTGTGCGCG 5100
5347 ACCGGGTGTGTGATCATATCGGCGGCTTGCATTAACAGAGAGCGCTGTGCGCG 5406
5101 GACAGAGAGTCTCTATAGAGGCTTTTATGAGATGAGAGAAATGTGCTTGAAGCGGCT 5160
5407 GACAGAGAGTCTCTATAGAGGCTTTTATGAGATGAGAGAAATGTGCTTGAAGCGGCT 5466
5161 CTCTTTGAGAGAGGGGACGCGGATATGCGGATGTGTAAGTCAAGATTCAGAGCTTAT 5220
5467 CTCTTTGAGAGAGGGGACGCGGATATGCGGATGTGTAAGTCAAGATTCAGAGCTTAT 5526
5221 CAGCAAGCTCTCAAAAGCTCAAGACATCAACCCCATGAGAGCTTATGAGCGGCA 5280
5527 CAGCAAGCTCTCAAAAGCTCAAGACATCAACCCCATGAGAGCTTATGAGCGGCA 5586
5281 GTAGAACAAATTTGAGGCGCAACATGTGAGACTTATAGCGGATCAATACCTGCA 5340
5587 GTAGAACAAATTTGAGGCGCAACATGTGAGACTTATAGCGGATCAATACCTGCA 5646
5341 GGAATTAACAATCTGCCAGGAAACCTGCAATGCTTCAATGATGGCTTCAATGCTGCC 5400
5647 GGAATTAACAATCTGCCAGGAAACCTGCAATGCTTCAATGATGGCTTCAATGCTGCC 5706
5401 CTCACAGTCCGCGTCAACAGAACCACTATCTTCTCAACATTTTGGGGGCTGACTA 5460
5707 CTCACAGTCCGCGTCAACAGAACCACTATCTTCTCAACATTTTGGGGGCTGACTA 5766
5461 GCATCCCAAAATTCACCAACCGCGGGGCACTGAGCTTCTGTGCAATGAGCTTATGAG 5520
5767 GCATCCCAAAATTCACCAACCGCGGGGCACTGAGCTTCTGTGCAATGAGCTTATGAG 5826
5521 GCTGCGTGAAGCAATAGGCTTATGATGATGTCTATGAGCAATCTGAGAGGATAGT 5580
5827 GCACTGTAGAGCAATAGGCTTATGATGATGTCTATGAGCAATCTGAGAGGATAGT 5886
5581 GCGGGCAATTCGAGGGGCTCTGCTGCAATTAAGATCAATGCTGAGAGAGGCTTCA 5640
5887 GCGGGCAATTCGAGGGGCTCTGCTGCAATTAAGATCAATGCTGAGAGAGGCTTCA 5946
5641 GAGAGTGTCTCAACTTGTGCTGCTGCAATTCGCTTCTGCGGATGCTTGTATGAGAGT 5700
5947 GAGAGTGTCTCAACTTGTGCTGCTGCAATTCGCTTCTGCGGATGCTTGTATGAGAGT 6006
5701 ATCTGCGCGGCTTCTGCGCGGCAAGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAG 5760
6007 ATCTGCGCGGCTTCTGCGCGGCAAGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAG 6066
5761 AATGACTCAATTCCTTGTCTTCAAGAGAAATCAGCTGCGCCCAACCACTAGTAC 5820

Db 6067 AACAGGCTTATGCTTGTGCTTCAGAGAAACACGCTGCCCCCACTACGTAGC 6126
Qy 5821 GAGTCGGATGCTGCGAGCGGTGACCCAACTACTTGGCTCCCTTACATTAACAGCGT 5880
Db 6127 GAGTCGGATGCTGCGAGCGGTGACCCAACTACTTGGCTCCCTTACATTAACAGCTA 6186
Qy 5881 CTCGAAGACTCCAACTGATTAAGAGACTGCCCCATCCCATCGGCGGCTGCTG 5940
Db 6187 CTCGAAGACTCCAACTGATTAAGAGACTGCCCCATCCCATCGGCGGCTGCTG 6246
Qy 5941 CTCGCGATGCTGCGAGCTGCGTTGACCACTCTTAACAGACTTTAAAAATTGGCTGAC 6000
Db 6247 CTCGCGACGCTGCGAGCTGCGTTGACCACTCTTAACAGACTTTAAAAATTGGCTGAC 6306
Qy 6001 TCCAAATTTATCCCAAGATGCGCGGCTCCCTTGTCTCTGCTCAAAAAGGGGTACAG 6060
Db 6307 TCCAAATTTATCCCAAGATGCGCGGCTCCCTTGTCTCTGCTCAAAAAGGGGTACAG 6366
Qy 6061 GGCCTGTGGGCGCGGCTGCGATGACCAACAGCTGTCTTGGCGGCGCAATATCT 6120
Db 6367 GGTGTCTGGGCTGCGATGATGATGACCAACAGCTGTCTTGGCGGCGCAATATCT 6426
Qy 6121 GGCATATGCTGCGCTTGGGCTCAAGAAATCAAGGGGCTTAAGACTTGAATATCTG 6180
Db 6427 GGTATGTCCGCTGGGCTCTATGAAATTAAGAGGCTTAAGACTTGAATATCTG 6486
Qy 6481 CAGGGGACTTTCCTATCAATTTGTTACAGAGGGGCGAGTGCCTGCAAAAAGGGGTAC 6240
Db 6487 CAGGGGACTTTCCTATCAATTTGTTACAGAGGGGCGAGTGCCTGCAAAAAGGGGTAC 6546
Qy 6241 AACTTTAAGTCCGCACTGAGAGGGTGGCGGCTCAGAGTACGCGAGAGTACGAGCAC 6300
Db 6547 AACTTTAAGTCCGCACTGAGAGGGTGGCGGCTCAGAGTACGCGAGAGTACGAGCAC 6606
Qy 6301 GGGTCATACCACTACATACAGAGACTCAACCTATTAAGTAAAGTCCCTGCACTA 6360
Db 6607 GGGTCATACCTCTATATAACAGAGTGAACCACTACATGGAAGTTCCTGCACTA 6666
Qy 6361 CCCCTCCGAGTCTTTTCTGAGTGAACGAGTGCAGATCCATAGTTTCCGCCACA 6420
Db 6667 CTTCTCAGAGTTTCTCTGAGTGAACGAGTGCAGATCCATAGTTTCCGCCACA 6726
Qy 6421 CCGAAGCGTTTTCGAGAGTGAAGTCTGTTCTGAGTGGGCTTAAATTGTTGTC 6480
Db 6727 CCAAGCGGTTTTTCCGAGTGAAGTCTGTTCTGAGTGGGCTTAAATTGTTGTC 6786
Qy 6481 GGGTCCAGCTTCTTGGAGCCCTGAACCCGACACAGCTATGATGTCCATGCTAAC 6540
Db 6787 GGGTCCAGCTTCTTGGAGCCCTGAACCCGACACAGCTATGATGTCCATGCTAAC 6846
Qy 6541 GATCATCTCATATCAAGCGAGAGACTGCGAGCGGCTTAAAGCGGGGTCACCCCA 6600
Db 6847 GATCATCTCATATTAAGCGAGAGCGCGAGCGGCTTAAAGCGGGGTCACCCCA 6906
Qy 6601 TCCGAGCGAGCTCTCGGCGAGCCAGCTATCGGCAACATGCTGCGAGCCACTGAC 6660
Db 6907 TCCGAGCGAGCTCTCGGCGAGCCAGCTATCGGCAACATGCTGCGAGCCACTGAC 6966
Qy 6661 ACCCAAGCGAAGCTCTATGATGTGACATGTTGATGCTAACTGTTCAAGGGGGGAT 6720
Db 6967 ACCCAAGCGAAGCTCTATGATGTGACATGTTGATGCTAACTGTTCAAGGGGGGAT 7026
Qy 6721 GTGACTCGATAGAGTCTGGTCCAAAGTGTGCTTCTGACTCTCTGACCAATGTC 6780
Db 7027 GTGACTCGATAGAGTCTGGTCCAAAGTGTGCTTCTGACTCTCTGACCAATGTC 7086
Qy 6781 GAAGAAAGAGAGCTTTGAGCTTCAATACATCAAGATACATGCTTCCCAAGAGAG 6840
Db 7087 GAAGAAAGAGAGCTTTGAGCTTCAATACATCAAGATACATGCTTCCCAAGAGAG 7146
Qy 6841 TTCCACCAAGCTTAAACCGGCTGGGCAAGGCTGATTAACCAACCGCTTGTGAATCG 6900

Db 7147 TTCCCGCAGCCTTACCGGCTGGGCAAGGCTGAGTACCAACCGCTTGTGAACCG 7206
Qy 6901 TGGAAAAAGGCAAGATTACCAACCGGCACTGTTGCGGGCTGTGCTCTCTCTCTAG 6960
Db 7207 TGGAAAAAGGCAAGATTACCAACCGGCACTGTTGCGGGCTGTGCTCTCTCTCTAG 7266
Qy 6961 AAAAAGGCAAGCTTCCCAAGAGAGCGCGGCAAGTGGGCTTAAAGTGAAGCTTCATA 7020
Db 7267 AAAAAGGCAAGCTTCCCAAGAGAGCGCGGCAAGTGGGCTTAAAGTGAAGCTTCATA 7326
Qy 7021 GAGATAGCTTCAAGAGTGCATTAAGTCTTGGGCAAGGCTTCAAGGCGGCTGAT 7080
Db 7327 GAGATAGCTTCAAGAGTGCATTAAGTCTTGGGCAAGGCTTCAAGGCGGCTGAT 7386
Qy 7081 TCAGGCTTTCAGAGGAGGCGGCTGCGATTCGCGAGTCAAGGCTCTGATGAG 7140
Db 7387 TCAGGCTTTCAGAGGAGGCGGCTGCGATTCGCGAGTCAAGGCTCTGATGAG 7446
Qy 7141 TTGGCTTTCAGAGAGAGGTTTCATCTCTTCAATGCTCCCTCGAGGGGAGCTTGA 7200
Db 7447 TTGGCTTTCAGAGAGAGGTTTCATCTCTTCAATGCTCCCTCGAGGGGAGCTTGA 7506
Qy 7201 GATCCAGACTGAGAGCTGAGAGTGAAGGCGGCAAGGCTTCAAGGCTTGAAG 7260
Db 7507 GATCCAGACTGAGAGCTGAGAGTGAAGGCGGCTTCAAGGCTTGAAG 7566
Qy 7261 GCTCCGAGCTGAGAGCTGAGAGTGAAGGCGGCTTCAAGGCTTGAAG 7320
Db 7567 GCTCCGAGCTGAGAGCTGAGAGTGAAGGCGGCTTCAAGGCTTGAAG 7626
Qy 7321 TGTCTCTCAATGCTATGCTGAGAGCGGAGCTTAAATCTCTTGTAGTCCGAAG 7380
Db 7627 TGTCTCTCAATGCTATGCTGAGAGCGGAGCTTAAATCTCTTGTAGTCCGAAG 7686
Qy 7381 GAGAATTAACCAATTAAGCTTGAAGCACTCCCTGCTGATGATCAACCAAGGCTG 7440
Db 7687 GAGAATTAACCAATTAAGCTTGAAGCACTCCCTGCTGATGATCAACCAAGGCTG 7746
Qy 7441 TGTACCAACCAAGAGCGCTCTCAAGAGGCTTGAAGGCTTGAAGGCTTGAAG 7500
Db 7747 TGTACCAACCAAGAGCGCTCTCAAGAGGCTTGAAGGCTTGAAGGCTTGAAG 7806
Qy 7501 GTGCTGACTCTATCAAGCTCAAGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 7560
Db 7807 GTGCTGACTCTATCAAGCTCAAGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 7866
Qy 7561 ACCGCAAGGCTCTCAAGCTCAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 7620
Db 7867 ACCGCAAGGCTCTCAAGCTCTCAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 7926
Qy 7621 TCTAAATATGAGTTTGGGCTAAAGAGTCCGAGCTTGTCCGAGAGGCGCTTAA 7680
Db 7927 TCTAAATATGAGTTTGGGCTAAAGAGTCCGAGCTTGTCCGAGAGGCGCTTAA 7986
Qy 7681 ATCAAGTCCGTGGAAGAGCTCTGGAAGAGCTCAACCAATCTTCAAGGCTTGA 7740
Db 7987 ATCAAGTCCGTGGAAGAGCTCTGGAAGAGCTCAACCAATCTTCAAGGCTTGA 8046
Qy 7741 ATGGCAAAATGAGGTTGTTGCGTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGA 7800
Db 8047 ATGGCAAAATGAGGTTGTTGCGTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGA 8106
Qy 7801 CTATGCTTACCTGAGCTGAGGCTGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGA 7860
Db 8107 CTATGCTTACCTGAGCTGAGGCTGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGA 8166
Qy 7861 ACACAAAATCTTCTCAAGGCTGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 7920
Db 8167 ACACAAAATCTTCTCAAGGCTGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 8226
Qy 7921 CAGCGGCTGAGGTTCTCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 7980
Db 8227 CAGCGGCTGAGGTTCTCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 8286

OY	7981	TATGATACCGCATGGTTTGTAAGTCAACCGCTCACTGAGAGAGACATCAAGACCTGAGAGATCC	8040
Db	8287	TATGATACCGCATGGTTTGTAAGTCAACCGCTCACTGAGAGAGATCAAGACCTGAGAGATCC	8346
OY	8041	ATATATCGGGGCTGCTCTCTTGGCCCGAGAGAGGCCACATGCGCATACACTCGCTAACTGAG	8100
Db	8347	ATATATCTGGCGCTGCTCTCTTGGCCCGAGAGAGGCCACATGCGCATACACTCGCTAACTGAG	8406
OY	8101	AGACTTTTACGTGGAGAGGGGCTATATGTTCAACAGCAAGAGGGCCAAACCTGGGGGTACAGGCGT	8160
Db	8407	AGACTTTTACGTGGAGAGGGGCTATATGTTCAACAGCAAGAGGGCCAAACCTGGGGGTACAGGCGT	8466
OY	8161	TGCCGCGCCACCGGGGCTCTCAACCACTGAGCTGGGGAAACCATATCAATGCTGAGTAA	8220
Db	8467	TGTGCGCGCCACCGGGGCTCTCAACCACTGAGCTGGGGTAAACCATATGCTGAGTAA	8526
OY	8221	GCCTTTAGCGGCTTGTAAAGCTTGCAAGGATTAATCGCGCCACAAATGCTGTATGCGGCGAT	8280
Db	8527	GCCCTTAGCGGCTTGTGAAGGCGCGGGGATTAATGTTCAACCAATGCTGTATGCGGCGAC	8586
OY	8281	GACTTGTGTTCATCTCTGAAAGCCAGGGGACCGAGAGGACAGACGGAAACCTBAGAGCC	8340
Db	8587	GACTTGTGTTCATCTCTGAAAGCCAGGGGACCGAGAGGACAGACGGAAACCTBAGAGCC	8646
OY	8341	TTCAACGAGGCTATGACCAAGGTAATTTCTGCCCTCTGTGTGACCCCCCAGACCGGAGTAT	8400
Db	8647	TTCAACGAGGCTATGACCAAGGTAATTTCTGCCCTCTGTGTGACCCCCCAGACCGGAAATAT	8706
OY	8401	GATCTGAGAGCTGATTAACATCTTTGCTCTCTCAAAATGTGTCTGGGCGCTGGGCCCAAGGC	8460
Db	8707	GACCTGGAGCTGATTAACATCTTTGCTCTCTCAAAATGTGTCTGGGCACTCGGCCCAAGGGC	8766
OY	8461	CGCCGCAATTAACCTACCTGACACAGAGACCCTTACCACTCCAAATCGGCCGGGCTCGCTGGGAA	8520
Db	8767	CGCCGCAATTAACCTACCTGACACAGAGACCCTTACCACTCCAAATCGGCCGGGCTCGCTGGGAA	8826
OY	8521	ACAGTTAGACACTCCCTGTGCAATTCATGAGCTGGGAAACATCATCCAGTACGCCCGAC	8580
Db	8827	ACAGTTAGACACTCCCTGTGCAATTCATGAGCTGGGAAACATCATCCAGTACGCCCTCAACC	8886
OY	8581	ATTATGGGTCCGCAATGTCCTGATGACACACTTTCTTCTTCATTTCTCATATGCTCAAGACAG	8640
Db	8887	ATTATGGGTTCGAATATGTCCTGATGACACACTTTCTTCTTCATTTCTCATATGCTCCCAAGACACT	8946
OY	8641	CTGGAACCAAGACTCAACTTGTAGATGTAACGAGACGGGTACTCGTGAATGCCCTTGGAC	8700
Db	8947	CTGGAACCAAGACTTTAACCTTGAATATGACGGAATCGGTATATTCGATBAGCCCTTTGGAC	9006
OY	8701	CTCCCAAGCTAATTTGAAAGGTTACATGAGGCTTGAACGCTTTTCTCTGCAACATACACT	8760
Db	9007	CTCCCAAGCTAATTTGAAAGGCTTACACAGGCTTGAACGCTTTCTCTCTGCAACATACACT	9066
OY	8761	CCCCACGAACCTGACACACGGGTGGCTTACGCCCTCAAGAAACCTTGGGGGGCCACCCCTCAGA	8820
Db	9067	CCCCACGAACCTTACACGGGTGGCTTACGCCCTCAAGAAACCTTGGGGGGCCACCCCTCAGG	9126
OY	8821	GCGTGAAGAGACCGGGCACTGACAGTCAAGGCGTCCCTCATCTCCGTGGGGGAGAGCG	8880
Db	9127	ACGTGAAGAGATCGGGCGCGTGAAGTTAGGGCGTCCCTCATCTCCATGGGGGGAGAGCG	9186
OY	8881	GCCGTTTTCGGGTCCATATCTCTTCAATTGGGGCGGTGAAGACCAAGCTCAAACTCACTTCA	8940
Db	9187	GCCATTTTCGGGTCCATATCTCTTCAACGTGGGCGGTGAAGAACCAAGCTCAAACTCACTTCA	9246
OY	8941	TTGGCGGAAGGCGCTCTCGATTTTATCCAGCTGAGTTCAACCGTCCGCGCGCGGGGGGGC	9000
Db	9247	TTGGCGGAAGGCGCTCTCGATTTTATCCAGTTGGTTTACGATCGGCGCGGGGGGGC	9306
OY	9001	GACATTTTATCACAGGCTGTCCGCTGGGCCCGACCCCGCTTATGTGCTTTTGGCTTACTCTTA	9060
Db	9307	GACATTTTATCACAGGCTGTCCGCTGGGCCCGACCCCGCTTATGTGCTTATGTGCTTACTCTTA	9366

QY	9061	CTTTTGTAGAGGGTAGAGCCTTTCTCTACTCCCGCTCG	9098
Db	9367	CTTTCGGTAGGAGGTAGCCTCTTCTACTCCCCCTCG	9404
RESULT 11	AF169004	9653 bp RNA linear	VRL 06-SEP-2001
LOCUS	Hepatitis C virus isolate G2ak3	polyprotein gene, complete cds.	
DEFINITION	AF169004		
ACCESSION	AF169004.1	GI:6707283	
VERSION			
KEYWORDS			
SOURCE	Hepatitis C virus		
ORGANISM	Hepatitis C virus		
REFERENCE	1 (bases 1 to 9653)		
AUTHORS	Kurihara, C., Ishiyama, N., Nishiyama, Y., Fukushi, S., Kageyama, T., Katayama, K. and Miura, S.		
TITLE	Molecular characterization of hepatitis C virus genotype 2a from the entire sequences of four isolates		
JOURNAL	J. Med. Virol. 64 (4), 466-475 (2001)		
MEDLINE	21561470		
LOCUS	11468731		
REFERENCE	2 (bases 1 to 9653)		
AUTHORS	Katayama, K. and Kurihara, C.		
TITLE	Direct Submission		
JOURNAL	Submitted (14-JUL-1999) Basic Research Division, Biomedical Laboratories, Inc., 1361-1 Metcoba, Kawagoe, Saitama 350-1101, Japan		
FEATURES	Location/Qualifiers		
SOURCE	1..9653		
	/organism="Hepatitis C virus"		
	/mol_type="genomic RNA"		
	/isolate="G2ak3"		
	/db_xref="taxon:11103"		
	/note="genotype: 2a"		
	341..9442		
	/codon_start=1		
	/product="polyprotein"		
	/protein_id="AAP25612.1"		
	/db_xref="GI:6707284"		
	/translation="MSTKPKPKQKTKRNTRRPDQYKFPFGGQIVGGVYLIPRRGRLGVATKRTSRSOPGRKROPIDRDRSTGSKWGPYWPVLVYEGELGAGWMLSPRGSRBWPQTPDRHRSRNVGKXIDITLTCFADLMQIIPVAGPLGVAALAHGVLEADGVNATDILPGCSFSLFLLALSLCITIPSAAYKNTSTYMTNDCSNIITQWLOLAVLPGCVCEKVGKXKSRMIPSPVPAVAVORGALITGRTHIDIMVMAATLSALYVGLCGVMLEAQMPTFISPOHMFVQECNITSPITGTHRMAMMMNSPATYMLAYAMRVEVYIIDITGGMHVMGRLATFNSQGMAYKVVYLLLAAGDYATXTRYAGKXAGVSPGIXGLFKPGAKONKQILINSNGWHINRTALNDSLELNGFLASLYTHSPNSGCEPRSLASCRNIEAFRIGWGTQIQEENVTNEDMRPYCMYPRQCGIVAESVCVSYCYTFPVPVVVGTTRDLGVPPYTWGENETDVFLENTFSPGSGPWGTMMNSGTTRKACGAPPCRRADPFNASTDLLCPDCEFRKHEPATVKKSGSPWLTPCLVDYPIRLMHPCTVNTYTIKIMVYGVGVHRLTAACNFSRSGRCLLEDREDOVSLPSTHTEMMALPCTSYDLPALSTGLILHONIVDOVYMGSLPALTXYVVMREVVYLLFLLAAPVACACIMMLILLEGAEALAEKLVYLLHAASAASCNGFLPIYPIPTAANYIGKRYPLATYSLTGLMPCFLLALPQOAVAYDGSVHGQIGALLITLFTLTPYAVKTLTSLCMMWLCYLLTLGEMVOWEAPSMQARGRDI IWAATIPCGVGVETITKMLAVLGPVYLLRDLATRPVYFAHALLMTCYVRLHLAGRYVQVMAALAGMTGYIYDHLTPMSDMAASGLRDILAAVEPIIFSPMEKKVITWGAETACGILHGLPVSATLGEIILGPADYASGLRDILAAITAYVNOOTRGLLGAIVSMGTGRDKTEQAGEIOLVSTVQSFLGTSIXVLMTVYHGAAGTKTAGSRGGLVQVMSAEGDVGSPSPGKSLSEPCTCGAVDLYITRNADVIPAARRDKRGALLSRPLSTLKGSSGPPVLCRGLAVGIFRAVCSRGYAKSIDTIPVEALATVTRSPFSDNSIPPAVPOTYOVGLYHPTSGSKSTKVPAVYAQGVYVVLNVPVLAATLPGFAYLSRAHGINPNIRTVRTVTGAATYISYGFADGGCGXGAVYI IICD ECHGSVSTILLIGTVLDOAETAGVRLTVLATPSPGVTTPPHNIEVALGQGEIIPFYGAGPLASYIKGGRHLIFCHSKKCDLAAALRGMGLNAVYVRGLDAVYI PTOGDVYVVAATDLMGTGYGDPSTVDCAVAVTVOVPSLDPEFTITTOTVODDARSQORRERTGGGRGIVYVSTGGRASGMSPSVLCEVDGAAAMELTPIVETTYRLAAYNTNCELPCVDHLEFWEAVFTGTHIDAFILSOTROSGENFAYLVAYQATVCARAKAPPSPNDIMMKLTVRLFTLVPFTLRLGSVTNEVYLLPVTXYLATFQMEADEASVATLTSWVLGVAALVAAYCLATGCVSIIIGRLHAXNOOVNVAIPKEVLYEFTDMEDECASALALEEQRIEMBLXSKIOGLLOOASKOADIOPVAVMSRVRVOPFAKHMMNFISGIVQIAGLSTGLPQVAMLSKIOGLLOOASPLSTSTILLINLGCMGLASQIAPRAGATGFPVSGVGVNAA		

VGIIGLKVLDVILAGVAGISGLVAFKINSKPSMEDVNLPGILSPGLVGV
ICAILRHHVPGEGAQOMNRLLAFASRGNHVPATRYVESDSQVTDLSLIT
SLRLRHHMITEDEPIPCASGMLRDVMDVMSILIDFNWMLSLPKMGLPLISQC
KYGKVMAGTGMTTRCPGANISGNVSGMSYITGTLTDNCKVPOULSPSPFMSV
VKBAPNFKTAIWRVAASEVAETOHGVSYSITGTLTDNCKVPOULSPSPFMSV
OIHRAFPTRKPFPRFDEVSFCVGLASPVVGSOLPCDPRPDVILTSLTDSHTAEEA
AKRLARSPSEBASASQULASPLRATCTTGKTYDMDADALFNGSDVTRJESBS
KIVLSDSDSWVEGSDHPSIPSEYMLPRTPPALPVMARPNPVPVSWRDPY
QPTVAGCALPPSKETPPPRRRRTVGLSESTIGDLQOALITFCQPPAPVSGSKS
TGADADSGRTPELTPSETGSTSSMPLEGBDPELPEVLEPOPEXVAP
GDSGMSITSEEDSVVCCMSYSMTGALITPCSPEBKPIPNXNSILRYHNKY
CTSKASLRAKYTPEDRMVLDHYSLDITGLASKYSARLITBEACQLTPHS
AKSKTGPASKEVSLSRBAVNHISVMKDLJEDSGTPTTITAKNVEFCPEPTGSK
KAARLIVPDLGVCEKRMALYDVTOKLPQAVMGASGFQSPORVEFLIKAWADK
DPMGSDYTRCFDSVTERDIRTESIYOACSLPEBARTAIHSILTERLYOGPWFNSK
GOTCGYRCRASGVLTSMGNTITCYVALAACAKAGVAPVATVIVIESGQ
ABEDBRNLRFVTEAMTRYASAPSDPREVDLELITGSSVSVASRPODRRYVILTR
DPTPIAAMETVPHFPVNSWMLGNIIOYAPTIVRWYLMHFPISIIAODTLDONIN
PMTGSVTSVSPDIPATIRHGLADASLHTVPHETRVASLARKGAPLHAMKS
RARAVRABLISHGGAACVGRYLFNMAVXKLKPLPEARLIDLSSMFTVAGAGGDI
VLCVLRARPRILLLSLILSVGLIFLIPAR"

ORIGIN

Query Match 86.8%; Score 7893.6; DB 14; Length 9653;
Beet Local Similarity 91.4%; Pred. No. 0;
Matches 8317; Conservative 39; Mismatches 742; Indels 0; Gaps 0;

1 ATGACACAAATCTTAAACCTCAAGAAAACCAAAAGAAACCAACCGTCGCGCAAA 60
341 ATGACACAAATCTTAAACCTCAAGAAAACCAAAAGAAATCTAACCGTCGCGCAAA 400
61 GACGTTAAGTTCCCGGGCGCGCGCAGATCGTTGGCGAGTATACCTTGTCGCGCAGG 120
401 GACGTTAAGTTCCCGGGCGCGCGCAGATCGTTGGCGAGTATACCTTGTCGCGCAGG 460
121 GGGCCAGGTTGGGTGTGCGCGCAGCAAGAAAGACTTGGAGCGGTCCCAACCACTGGA 180
461 GGGCCAGGTTGGGTGTGCGCGCAGCAAGAAAGACTTGGAGCGGTCCCAACCACTGGA 520
181 AGGCGCAGCCCATCTTAAAGATCGGCGCTCACTGGCAATCTCTGGGAAAACCAAGA 240
521 AGGCGCAGCCCATCTTAAAGATCGGCGCTCACTGGCAAGTCTCTGGGAAAACCAAGA 580
241 TACCCCTGGCCCTTATACGGAATGAGGGAATCGGCTGGGCAAGATGCTCTGTCCTCC 300
581 TACCCCTGGCCCTTATATGGAATGAGGGAATCGGCTGGGCAAGATGCTCTGTCCTCC 640
301 CGAGGTTCCCGTCTCTTGGGGCCCAATGACCCCGGCAATAGTCCGCGCAAGTGGGT 360
641 CGAGGTTCCCGTCTCTTGGGGCCCAATGACCCCGGCAATAGTCCGCGCAAGTGGGT 700
361 AAGGTCATGATACCTTAAAGTGGCGCTTGGCGACTCATGGGGTATACATCCCTGTGCTG 420
701 AAGGTCATGATACCTTAAAGTGGCGCTTGGCGACTCATGGGGTATACATCCCTGTGCTG 760
421 GGGCGCCCGCTGGGCGCGCTCGCAGAGCTCTCGGCAATGGCGTGAAGTCTCTGAGGAC 480
761 GGGCGCCCGCTGGGCGCGCTCGCAGAGCTCTCGGCAATGGCGTGAAGTCTCTGAGGAC 820
481 GGGGTTAATTTTGGCAACAGGGAATTTACCGGTTGCTCTTTCTATCTTCTGTGGCC 540
821 GGGGTTAATTTTGGCAACAGGGAATTTACCGGTTGCTCTTTCTATCTTCTGTGGCC 880
541 CTGCTGTCTGTGATCAACACCCCGGCTCTCGGCTGCGAAGTGAAGATCATGATCGGCG 600
881 CTACTGTCTGTGATCAACATCCCGGCTCTGTGCTGCGCAGGTGAAGAACACAGCACCGGT 940
601 TACATGTGTGATCAACAGTGAACCAATGACAGACTTACCTGGCAGCTCAAGGCTGTGTC 660
941 TACATGTGTGATCAACAGTGTGCAATGACAGACTTACCTGGCAGCTCAAGGCTGTGTC 1000
661 CTCACGTGTGTGCGGTGCGTCCCGTGGCAAGAGTGGGAATGATCTTCAGTGTGATTA 720

1001 CTCACGTCCTCCGCGGTGTCTCCGTTGTGAGAAAGTGGGAATATATATACGCTGTGGATA 1060
721 CCGGTCTACCGGAATGTGGCCGCTGACGAGCGCCGCGCCTGACGAGGGCTTGGCGAGC 780
1061 CCGGTCTACCGGAATGTGGCCGCTGACGAGCGCCGCGCCTGACGAGGGCTTGGCGAGC 1120
781 CACATGACATGTTGTGATGTCTCCGCAAGCTCTGCTCTGCTCTTACGTGGGGACCTC 840
1121 CACATGACATGTTGTGATGTCTCCGCAAGCTCTGCTCTGCTCTTACGTGGGGACCTC 1180
841 TGGGTGGGGTGTGATGTCTCCGCAAGCTCTGCTCTGCTCTTACGTGGGGACCTC 900
1181 TGGGTGGGGTGTGATGTCTCCGCAAGCTCTGCTCTGCTCTTACGTGGGGACCTC 1240
901 GTCGAAGATGCAATGCTCTCACTTCACTCGGTGATACCATCACTGACACCGCATGGCATG 960
1241 GTGCAAGATGCAATGCTCTCACTTCACTCGGTGATACCATCACTGACACCGCATGGCATG 1300
961 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
1301 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1360
1021 CCGGAGTGAATTTAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
1361 CCGGAGTGAATTTAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1420
1081 TACTTCTTATGACGAGGAGCGTGGCGAAAGTGTGTGATCTCTTCTGTTGGCGCGCGG 1140
1421 TACTTCTTATGACGAGGAGCGTGGCGAAAGTGTGTGATCTCTTCTGTTGGCGCGCGG 1480
1141 GTGACGCGCGCACCATATCTGTTGGGGTCTGCGCGCAGACACCGGGCGCTCAC 1200
1481 GTGACGCGCGCACCATATCTGTTGGGGTCTGCGCGCAGACACCGGGCGCTCAC 1540
1201 AGCTTATTTGACATGAGGCGCCGACGAGAAATACAGCTGTGTTAACAACCAATGGCAGCTG 1260
1541 AGCTTATTTGACATGAGGCGCCGACGAGAAATACAGCTGTGTTAACAACCAATGGCAGCTG 1600
1261 CACATCAACCGCACCGCCCTGAACTGCAATGATCTCTTGCACACCGGCTTATCGCGTCT 1320
1601 CACATCAACCGCACCGCCCTGAACTGCAATGATCTCTTGCACACCGGCTTATCGCGTCT 1660
1321 CTGTTCTACACCAAGCTTCACTGTGACAGATGTCGGAACGATGTCCTGCTGCGC 1380
1661 CTGTTCTACACCAAGCTTCACTGTGACAGATGTCGGAACGATGTCCTGCTGCGC 1720
1381 AGTATGAGGCGCTTCCGGGTGGGATGGGGCGCTTGCATATGAGGATATATGACCAAT 1440
1721 AGTATGAGGCGCTTCCGGGTGGGATGGGGCGCTTGCATATGAGGATATATGACCAAT 1780
1441 CCAAGAGATATGAGACCTTATGCTGCACTTACCAACCAAGGAGTGTGGCTGTCTCC 1500
1781 CCAAGAGATATGAGACCTTATGCTGCACTTACCAACCAAGGAGTGTGGCTGTCTCC 1840
1501 GCGAAGACTGTGTGCGCCAGTGTACTGTTTCAACCCCAAGCCAGTGTGTGGGACAG 1560
1841 GCGAAGACTGTGTGCGCCAGTGTACTGTTTCAACCCCAAGCCAGTGTGTGGGACAG 1900
1561 ACCGACAGGCTTGAAGGCGCCACTTACAGTGGGGGGGAGATGAGACAGATGCTTCTTA 1620
1901 ACCGACAGGCTTGAAGGCGCCACTTACAGTGGGGGGGAGATGAGACAGATGCTTCTTA 1960
1621 TTGAACAGCACTGACACCGCTGAGGATCATGTTGCGCTGACAGTGTGAATCTTCTT 1680
1961 TTGAACAGCACTGACACCGCTGAGGATCATGTTGCGCTGACAGTGTGAATCTTCTT 2020
1681 GGCTACACCAAGCTTGGCGGACCAACCTGCGCTGATGAGACTGATCAACGCGCAGC 1740
2021 GGCTACACCAAGCTTGGCGGACCAACCTGCGCTGATGAGACTGATGAGACTGATGAGACT 2080
1741 ACCGACGCTGTTGGCGCCAGCACTGTTTGAAGAGATCTGATACCACTTCACTCAAA 1800
2081 ACCGACGCTGTTGGCGCCAGCACTGTTTGAAGAGATCTGATACCACTTCACTCAAA 2140

QY	1801	GGCGGCTCTGGGGCCCTGGCTCAGCGCAAGGCGCTGATCGACTACCTTCACAGGCTCTGG	1860
Db	2141	TGTGGTCTGGAACCCCTGGCTTAACGCCAAGGTCGTGGCTGACATACCCCTACAGGCTCTGG	2200
QY	1861	CATTACCCCTGACAGATTAACTATACACTCTTCAAAAATAGATGTATGTGGAGGGGTT	1920
Db	2201	CATTACCCCTGACAGATTAACTATACACTCTTCAAGAATAGATGTATGTGGGGGGGTT	2260
QY	1921	GAGACAGGCTCACAGGCTGCAATGCAATTTGCACTGTGGGGATGCTGTGCACATTGGAGAC	1980
Db	2261	GAGACAGGCTCACAGGCGCGCGGTGCAACTCTCTGTGGGTATGCGTGGCAATTGGAGAC	2320
QY	1981	AGAGACAGAAATCAACTGTCTCTTGTGTGACCTCACACAGAAATGGGCAATTTACT	2040
Db	2321	AGGACAGAGGTCAAAGTGTCTCTTGTGTGACCTCACACAGAGTGGGCAATTTGCCC	2380
QY	2041	TGCTCTTACTCGGACCTGCGCGCTGTGCACTGTCTTCTCACTTCACCAAAACATC	2100
Db	2381	TGCTCTTAACACAGACTGCGCTGTGTGCTCCACTGTCTTATTCACCTTCACCAAAATATC	2440
QY	2101	GTGAGACGAACTTATCATGTATGGCTATCACTGCGCCCAAAAATACATGTGCGAATGG	2160
Db	2441	GTGAGACGTACATACATGTATGGCTGTCACTGCGCCCAAAAATATGTGTGTGCGAATGG	2500
QY	2161	GAGTGGGTAAATACTTATATTCCTGCTCTTACCGGACGCGCAGAGTTTGGCGCTGTATGG	2220
Db	2501	GAGTGGGTAAATACTTATATTCCTGCTCTTACCGGACGCGCAGAGTTTGGCGCTGTATGG	2560
QY	2221	ATGCTCATCTTGTGGGCGAGGCGGAGACGACATAGAAAGCTGTATCTTGGACAGCT	2280
Db	2561	ATGCTCATCTTGTGGGCGCAAGCGGAGACGACATAGAAAGCTGTATCTTGGACAGCT	2620
QY	2281	GCGAGCGAGCTAGCTGCAATAGGCTCTCTAATTTTGTCACTCTTTGTCGTGGCTGTGG	2340
Db	2621	GCGAGTGAAGCTAGCTGCAATAGGCTCTCTAATCTTGTGTCACTCTTTGTCAGTGTGG	2680
QY	2341	TACATCAAGGTCGGGTATGTCCCTTAGACTAATTCCTCACTGCGCTGTGTCTTT	2400
Db	2681	TACATCAAGGTCGGGTATGTCCCTTAGACTAATTCCTCACTGCGCTGTGTGTCTTT	2740
QY	2401	AGCCTACCTGCTCCATAGCATTTGCCCAACAGGCTATGCTTATGACGATCTGTGCAATGGC	2460
Db	2741	TGCCTACCTGCTCCATAGCATTTGCCCAACAGGCTTATGCTTATGATGATCTGTGCAACGGA	2800
QY	2461	CAGATAGAGCGGCTGTGCTGTATGATCACTCTTACTCTCACCCCGGATATAG	2520
Db	2801	CAGATAGGCGGCGTGTGTGTATCTATTAACCTGTTCACATCAACCCCGGCTATATAG	2860
QY	2521	ACCTTCTCAGCCGGTTTTTGTGTGTGTGTGTCTATCTTTCGACCCCTGGGGGAAGCTATG	2580
Db	2861	ACCTTCTCAGCCGGTGTGTGTGTGTGTGTGTCTATCTTTCGACCCCTGGGGGAAGCATG	2920
QY	2581	GTCCAGAGATGGGACCACTATCAGGTGGCGCGGTGACGATCAATATATGGAGCC	2640
Db	2921	GTCCAGAGATGGGACCACTATCAGGTGGCGCGCGGTGACGATGATATATATGGAGCC	2980
QY	2641	GTCCCAATATTCTAACCAAGGTGTGTGTGTGTGACATTAACCAAGGTGCTCTTGGCGGTCTT	2700
Db	2981	GTCCCAATATTCTTCCCGGGGTGTGTGTGTGTGACATTAACCAAGGTGCTCTTGGCGGTCTT	3040
QY	2701	GGGCTGTGCTTACCTCTTAAAGGTGCTTGAACGGGCGTGGCCGTACTTGTGCAGGGCTCAC	2760
Db	3041	GGGCTGTGCTTATCTCTTAAAGATGCTTGTGAACGTGTGCCGTATTTGTGTGAAGCTCAC	3100
QY	2761	GCTCTACTAGAGATGTGACCAATGGCAAGCATCTGCGGGGGGCAAGTATGCTCAGATG	2820
Db	3101	GCTCTGTGAGAGATGTGACCAATGTGTAGGCACTCTGGAAGGGGTAGATGCTCAGATG	3160
QY	2821	GGGCTACTAGCCCTTGGCAGGTGCACTGCGCACTTAATCTATATGACACATCAACCCCTATG	2880
Db	3161	GGGCTACTATGAGCCCTTGGAGGTGCAAGGCACTTAATCTATATGATATCTCAACCCCTATG	3220

QY	2881	TTGGATTGGGCTGTAGTGGCTTCGGGGACTTCGGCGGTCCGGTGTAGACCTATCATCTTC	2940
Db	3221	TCGACATGGGCTGCGCATGGCTTCGGGACTTGGCGGTCGGCTGTGAGACTATCATCTTC	3280
QY	2941	AGTCGATGAGAAAGATCATTTGTCTGGGAGCGGAGACAGCTGCTTGTGGGACATT	3000
Db	3281	AGTCCGATGAGAAAGATCATCATCTGGGGAGCGGAGACGCGCTGATGGGGGAGATC	3340
QY	3001	TTTACACGACTTCCGTGTCCGCCGGCTAGAGCGGGAGATCTCTTGGCCACGTGAT	3060
Db	3341	TTTACACGACTTCTGTGTCCGCCGGCTAGAGCGGGAGATCTCTTGGCCACGTGAT	3400
QY	3061	GGCTATACCTTCGAAAGGGGTGGAGTCTTTCGCCCCCATCACTGCTTACCGCCACAGACA	3120
Db	3401	GGCTATACCTTCGAAAGGGGTGGAGGCTCTTGCCCCCATCACTGCTTACCGCCACAGACA	3460
QY	3121	CGTGGCCCTTTTGGGACCAATAGTGTAGCATGACGGGGCGCCGACAAAGACAGGCT	3180
Db	3461	CGAGGCGCTTTTGGGCGGCATAGTGTAGCATGACGGGGCGCCGACAAAGACAGAGCGCC	3520
QY	3181	GGGGAATTCGAGTCTGTGCCACAGTCACTCAGTCTTCCTTCGGAACATCATCTCGGG	3240
Db	3521	GGGGAATTCGAGTCTGTGTCTTACCGGTACATCAGTCTTCCTTCGGAACATCATCTCGGG	3580
QY	3241	GTTTGTGGACTGTCTACCATGAGCTGGCAACAGACTCTGCGCGGCTCACCGGGTCCG	3300
Db	3581	GTCCTGTGGAACCGGTCTACCATGAGCTGGCAACAGACTCTGAGCGGTTTACCGGGTCCA	3640
QY	3301	GTCACGAGATGTATCTCAGTGCAGAGGGGACCTTAGTAAAGTGGGCCACGCGCCCTGGG	3360
Db	3641	GTCACGAGATGTATCTGAGTGCAGAGGGAGACTTGTGTGGGTGGCCBAGCGCTCTGGG	3700
QY	3361	ACTTAATCTTTGAGCCGTGCACGTGTGAGCCGCTGCACCTGTATCCTGATCACCGCGAAC	3420
Db	3701	ACCAAGTCTTTGAGCCGTGCACGTGTGAGCCGCTGCACCTGTATCTGTATCAGCGAATC	3760
QY	3421	GCTATATGTCAATCCCGGCTCGAAGACGGGGGACAAACGGGGAGCGCTATCTCCCCGAGA	3480
Db	3761	GCTATATGTCAATCCCGGCTCGAAGACGGGGGACAAACGGGGAGCGCTATCTCCCCGAGA	3820
QY	3481	CCTCTTTCGACCTTGAAGGGGTCCTTCAGAGAGGCCCGGTGCTATATCCACAGGGCGCAGCT	3540
Db	3821	CCTCTTTCGACCTTGAAGGGGTCCTTCGAGGGGAGCCGCTGCTTTCCTTAGGGGCGACGCC	3880
QY	3541	GTCGGAATCTTTCGCGGAGCTGTGTGCTCTTCGCGGCGCTGAGTCAATAGATTTTCATC	3600
Db	3881	GTCGGAATCTTTCGCGGAGCTGTGTGCTCTTCGCGGCGCTGAGTCAATAGATTTTCATC	3940
QY	3601	CCCGTTGAGACATCTGCACATGTGTACGCGGATCCCCCATCTTTAGTGAACAACAGACACA	3660
Db	3941	CCCGTTGAGAGGCTGTGCATGTGTACGCGGATCTCCCATCTTTAGTGAACAACAGACACA	4000
QY	3661	CCTGCTGTGCCCCAGACTTATCAAGGTGGGATTACTTGCATGCCCGAATGGCAGTGAAGA	3720
Db	4001	CCGGCTGTGCCCCAGACTTATCAAGGTGGATTACTTGCATGCCCGAATGGCAGTGAAGA	4060
QY	3721	AGCACCAAAAGTCTGTGTGCATATGTCTGCACAGGGGTATTAAGTGTCTTAATCTCC	3780
Db	4061	AGTACCAAAAGTCTGTGTGCATATGTCTGCACAGGGGTATTAAGTGTCTTAATCTCT	4120
QY	3781	TCAATGTGCTGCACCTCGGGGTTTGGGGCTGTATTGTCTAAGGACATATGACATATCTCC	3840
Db	4121	TCTGTGCTGCACCTCAAGGTTTGGGGCTGTATTGTCTAAGAGCAATGGCATTAATCTCC	4180
QY	3841	AACATTAGACCTGAGATGACGACTGTGACACACCGGGGCGCCATCACTACTCCACTAT	3900
Db	4181	AACATTAGACCTGAGATGACAACTGTGACACACCGGGGCGCCATCACTACTCCACTAT	4240
QY	3901	GGCAAAATTCCTCGCCGATGGGGGCTGTGCGGGGCGCGCTTACACACATCATCATATGTGAT	3960
Db	4241	GGCAAAATTCCTCGCCGATGGGGGCTGTGCGGGGCGCGCTTATGACATCATCATATGTGAT	4300
QY	3961	GAATGCGATGCGGTGACTGTACCAACCATCTTGGGATCGAAACAGTCTTGTATCAAGCA	4020

Db 4301 GAAATGCCACTGTGGATTCTTACCAACATCTCGGCATCGGAACAGTCTTGAATCAAGCA 4360
Qy 4021 GAGACAGTGGGGTCAAGCTAACTGTGTGCTACAGTACGCCCCCTGGGTCTGTACA 4080
Db 4361 GAGACAGTGGGGTCAAGCTGTGTGCTACAGTACGCCCCCTGGGTCTGTACA 4420
Qy 4081 ACCCCCAACCCCAATAGAGAGGTGCCCCCTGGGAGAGAGGCGAGATCCCTTCTAT 4140
Db 4421 ACCCCCAATCCCAATATAGAGAGGTAGCCCTCGGAGAGAGGTGAGATCCCTTCTAT 4480
Qy 4141 GGGAGGGGAGTTCCTCTGTCTTATCATCAAGGAGAGAGATCTGATCTTCTGCATTTCA 4200
Db 4481 GGGAGGGGAGTTCCTCTGTCTTATCATCAAGGAGAGAGACTTGATCTTCTGCATTTCA 4540
Qy 4201 AAGAAAAAGTGAAGAGCTCGGCGGCGCCCTTCGGGGGTATGGGGTTGAACTCAGTGGCA 4260
Db 4541 AAGAAAAAGTGAAGAGCTCGGCGGCGCCCTTCGGGGGTATGGGGTTGAACTCAGTGGCA 4600
Qy 4261 TACTACAGAGGTTGGAGCTCTCGGTAAATACCAATCAAGGAGAGTATGATGCTGCC 4320
Db 4601 TATTACAGAGGTTGGAGCTCTCGGTAAATACCAATCAAGGAGAGTATGATGCTGCC 4660
Qy 4321 ACCGACGCCCTCATGACAGGGTATATCTGGGACCTTTGACTCCGTGATGACTGCAAGTA 4380
Db 4661 ACCGACGCCCTCATGACAGGGGTATATCTGGGACCTTTGACTCCGTGATGACTGCAAGTA 4720
Qy 4381 GGGGTCACTCAAGTTGTAGACTCTCAGTTTAAACCCCAATTCACATTAACACACAGATT 4440
Db 4721 GGGGTCACTCAAGTTGTAGACTCTCAGTTTAAACCCCAATTCACATTAACACACAGATT 4780
Qy 4441 GTCCCTCAAGAGCTGTCTCAAGTAGCAGCGCCGAGGTCGCACAGGATGAGGGAAGACTG 4500
Db 4781 GTCCCTCAAGAGCTGTCTCAAGTAGCAGCGCCGAGGTCGCACAGGATGAGGGAAGACTG 4840
Qy 4501 GGCATTTATAGTATGTTTCCACTGTGTAGGAGAGGCTTACAGAAATGTTTGAAGTGTAGT 4560
Db 4841 GGCATTTATAGTATGTTTCCACTGTGTAGGAGAGGCTTACAGAAATGTTTGAAGTGTAGT 4900
Qy 4561 CTCTGTAGTGTATGACAGAGGAGCGCATGTGTAGCTCAACCATTCGGAAGCACCC 4620
Db 4901 CTCTGTAGTGTATGACAGAGGAGCGCTGTGTGTATGAGCTCAACCATTCGGAAGCACCC 4960
Qy 4621 GTCAAGCTCAGGCGGTATTTCAACAGCGCCGGTTTCTGTGTGCAAGACATCTTAG 4680
Db 4961 GTCAAGCTCAGGCGGTATTTCAACAGCGCTTCAACAGCATAGATGCCCACTTCTTCCAAACA 5020
Qy 4681 TTTTGGAGAGGAGTTTTCACCGGGCTCACACATAGATGCCCACTTCTTCCAAACA 4740
Db 5021 TTTTGGAGAGGAGTTTTCACCGGGCTCACACATAGAGCGCCCACTTCTTCCAAACA 5080
Qy 4741 AAGCAATCGGGGGAATTTGCGATACTTAACAGGCTTACAGGCTTACAGTGTGCGCTAG 4800
Db 5081 AAGCAATCGGGGGAATTTGCGATACTTAACAGGCTTACAGGCTTACAGTGTGCGCTAG 5140
Qy 4801 GCCAAAGCCCCCCCCCGTCTGGAGCGTCAATGTGAAAGTGTGTTGACTCGACTCAAGCCC 4860
Db 5141 GCCAAAGCTCCCCCCCCCGTCTGGAGCAATCATGTGAAAGTGTGTTGACTCGACTCAAGCCC 5200
Qy 4861 ACACTGTGGGCCCCCAACACTTCTGTACCGCTTGGGCTGTGTTACCAAGAGTCAACC 4920
Db 5201 ACGCTGTGGGCCCCCAACACTTCTGTACCGCTTGGGCTGTGTTACCAAGAGTCAACC 5260
Qy 4921 CTCAACATCCCGTACGAAATATACGCACTGATGCAAGGAGGAGGAGGAGGCTGAGGTATG 4980
Db 5261 CTCAACATCCCGTACGAAATATACGCACTGATGCAAGGAGGAGGAGGAGGCTGAGGTATG 5320
Qy 4981 ACCAGACATAGGATCTTGGCAAGGAGGAGTCTTGGCGGCGTGCAGCGCTATTTGCTGCG 5040
Db 5321 ACCAGACATAGGATCTTGGCAAGGAGGAGTCTTGGCGGCGTGCAGCGCTATTTGCTGCG 5380
Qy 5041 ACCGGGTGTGTTGCAATCGGCGGCTTGCATTAACAGGAGGCGGTGTTGCGCG 5100

Db 5381 ACTGGGTGTGTTTCCATCATCGGCGCGGTTGACARTCAACAGCAAGTGTGTTGACCG 5440
Qy 5101 GACCAAGAGTCTCTATAGAGGCTTTGATGAGTGAAGAAATGTGCTCTAGGGCGGCT 5160
Db 5441 GACCAAGAGTCTCTTATAGGCTTTTACGAGATGAGAAATGTGCTCTAGAGCGCT 5500
Qy 5161 CTGATTGAAGAGGAGCGGATAGCGAGATGCTGAAGTCCAAAGCTTATTTG 5220
Db 5501 CTCCTTGAAGAGGAGCGGATAGCGAGATGCTGAAGTCCAAAGCTTATTTG 5560
Qy 5221 CAGCAAGTCTCCAAAGAGCTCAAGATGAAACCCAGTGCAGGCTTATGCGCCAG 5280
Db 5561 CAGCAAGTCTCCAAAGAGCTCAAGATGAAACCCAGTGCAGGCTTATGCGCCAG 5620
Qy 5281 GTAGAACAAATCTGGGCAACATGAGAACTTCATTAGGCGCATCCAAATCTGCA 5340
Db 5621 GTAGAACAAATCTGGGCAACATGAGAACTTCATTAGGCGCATCCAAATCTGCA 5680
Qy 5341 GGAATATCAACTGCGCAGGAAACCTGAGTAGTCTTCATGATGAGCTTCACTGCGCC 5400
Db 5681 GGAATATCAACTGCGCAGGAAACCTGAGTAGTCTTCATGATGAGCTTCACTGCGCC 5740
Qy 5401 CTCAACAGTCCGTGTCAACAGACACTATCTTCTCAATTTGGGGGCTGGCTTA 5460
Db 5741 CTCAACAGTCCGTGTCAACAGACACTATCTTCTCAATTTGGAGGCTGGCTTA 5800
Qy 5461 GCATCCCAATTTGACCAACCGCGGGGCACTGGCTTGTGTCAGTGGCTAGTGGGA 5520
Db 5801 GCATCCCAATTTGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 5860
Qy 5521 GCTGCGTAGAGAGTATAGGTTAGTAAAGTCTAGTGAACATCTGACAGGATAGT 5580
Db 5861 GCTGCGTAGAGAGTATAGGTTAGTAAAGTCTAGTGAACATCTGACAGGATAGT 5920
Qy 5581 GCGGGATTTTGGGGGCTCTCGTGCATTCAAGATATGCTGGGAGAGAGCCCTTCAG 5640
Db 5921 GCGGGATTTTGGGGGCTCTCGTGCATTCAAGATATGCTGGGAGAGAGCCCTTCAG 5980
Qy 5641 GAGAGTGTGTCAACTTGTGCTGACTGGAATTTCTCCGAGTCCCTGTGTAGTGGAGT 5700
Db 5981 GAGAGTGTGTCAACTTGTGCTGACTGGAATTTCTCCGAGTCCCTGTGTAGTGGAGT 6040
Qy 5701 ATCTGGGCGCATTTCTGGCGGACACGTTGGGACCGGGGAAAGGCGCTTCAATGGATG 5760
Db 6041 ATCTGGGCGCATTTCTGGCGGACACGTTGGGACCGGGGAAAGGCGCTTCAATGGATG 6100
Qy 5761 AATAGACTCATGSCCTTTGCTTCCAGAGGAAATCAAGTGGCGCCGACCCACTAACGAGAG 5820
Db 6101 AATAGACTCATGSCCTTTGCTTCCAGAGGAAATCAAGTGGCGCCGACCCACTAACGAGAG 6160
Qy 5821 GAGTGGATGCTGTGAGCGGTGACCCAACTACTTGAGCTCCCTTACATAACAGCTTG 5880
Db 6161 GAGTGGATGCTGTGAGCGGTGACCCAACTACTTGAGCTCCCTTACATAACAGCTTG 6220
Qy 5881 CTCAAGAGTCCCAACTGAGATTACTGAGAGCTGCCCATCCCATGCGGCGCTGTGG 5940
Db 6221 CTCAAGAGTCCCAACTGAGATTACTGAGAGCTGCCCATCCCATGCGGCGCTGTGG 6280
Qy 5941 CTCGGAGATGTGGGACGTTGGTTGACACATCTTAACAGACTTTAAATTTGCTGACC 6000
Db 6281 CTCGGAGATGTGGGACGTTGGTTGACACATCTTAACAGACTTTAAATTTGCTGACC 6340
Qy 6001 TCCAAATTTATCCCAAGATGCCGCGCTCCCTTTGTCTCTGTCAAAAAGGGTACAAG 6060
Db 6341 TCCAAATTTATCCCAAGATGCCGCGCTCCCTTTGTCTCTGTCAAAAAGGGTACAAG 6400
Qy 6061 GGGGTGTGGCGGCACTGGCATATGACCAACGGTGTCTTGGCGGCGCAATATCTCT 6120
Db 6401 GGGGTGTGGCGGCACTGGCATATGACCAACGGTGTCTTGGCGGCGCAATATCTCT 6460
Qy 6121 GCGAATCTCGGCTTGGCTTCAATGAGAAATCAAGGGGCTTAAGACCTGCAATGATCTGG 6180
Db 6461 GCGAATCTCGGCTTGGCTTCAATGAGAAATCAAGGGGCTTAAGACCTGCAATGATCTGG 6520

QY 6181 CAGGGGACCTTTCTATCAATTTGTTACAGGAGGGCCAGTGCCTGCCGAAACCCGGGCCA 6240
DB 6521 CATGGGACTTTTCTTATCAATTTGTTACAGGAGGGCCAGTGCCTGCCGAAACCCGGGCCA 6580
QY 6241 AACTTAAAGTCCGCACTTGAGAGGGTGGCGGCTCAGAGTACGGGAGTGAACGACAC 6300
DB 6581 AACTTAAAGACCGCACTTGAGAGGGTGGCGGCTCAGAGTACGGGAGTGAACGACAC 6640
QY 6301 GGGTCATACCACTTACATTAAGAGCTACCACTGATTAAGTAAAGTCCCTGCCTAACTA 6360
DB 6641 GGGTCATACCTTATATTAAGAGCTACCACTGATTAAGTAAAGTCCCTGCCTAACTA 6700
QY 6361 CCCCTCCCGAGTCTTTCTGAGGAGGAGGAGTGCAGATCCATAGGTTTGCCTCCACA 6420
DB 6701 CTTTCTCAGAGTTTCTTCTGAGTGAAGAGTGCAGATCCATAGGTTTGCCTCCACA 6760
QY 6421 CCGAAGCGTTTTCGCGGAGTGAAGTCTCGTTCTGCGTGGCTTAAATTCATTTGTCGC 6480
DB 6761 CCAAGCGTTTTCGCGGAGTGAAGTCTCGTTTTCGCTGGGCTCAATTCGTTGTCGC 6820
QY 6481 GGGTCCAGCTTCTTTCGAGCCTTGAACCCGACAGAGCTGATGTCATGCTTAA 6540
DB 6821 GGGTCTCAGCTTCTTGTGATCCAGAGCCGATACAGACGTAATGACGTCATGCTTAA 6880
QY 6541 GATCCATCTGATTAACAGGAGGAGCTGACAGCGCGGCTTAAAGCGGGGCTACCCCA 6600
DB 6881 GATCCATCCATTAACAGGAGGAGCTGACAGCGCGGCTTAAAGCGGGGCTACCCCA 6940
QY 6601 TCCGAGGCAAGCTCTTCGCGAGGAGCTTACGCACTGCTGCGAGCCACTGACC 6660
DB 6941 TCGAGGCGAGCTCTTCGCGAGGAGCTTACGCACTGCTGCGAGCCACTGACC 7000
QY 6661 ACCCAGCGCAAGCTTATGATGAGCATGAGTGAATCTTAACTGTTCAATGGGGGCGAT 6720
DB 7001 ACCCAGCGGAAAACTTATGATGAGCATGAGTGAATCTTAACTGTTCAATGGGGGCGAT 7060
QY 6721 GTGACTCGAGTAGAGTGTGGGTCCAAAGTGTGTTCTGACTCTCTCGACCCCAATGTC 6780
DB 7061 GTGACTCGAGTAGAGTGTGGGTCCAAAGTGTGTTCTGACTCTCTCGACCCCAATGTC 7120
QY 6781 GAAAGAAAGAGCAGCTTGAAGCTTGCATCAATCAGAAATATAGTCTCCCAAGAGAG 6840
DB 7121 GAAAGAGGAGCAGCAGAGCTTGCATCAATCAGAAATATAGTCTCCCAAGAGAG 7180
QY 6841 TTCCACAGCTTTACCGGCTTGAGGAGCGGCTGATTAACCCACCGCTTGTGAAATG 6900
DB 7181 TTCCCGCAGCTTGTGCGGTCTGGGAGCGGCTGATTAACCCACCGGTTGTGAAATG 7240
QY 6901 TGGAAAAAGCAGATTAACCAACCGGCACTGTCGCGGCTGTCTCTCCCTCCAG 6960
DB 7241 TGGAAAGGCGCAATTAACCAACCACTGTGCGGCTGTCTCTCCCTCCAG 7300
QY 6961 AAAACCCGAGCGCTCCCAAGAGAGGCGCGGAGCAGTGGGCTTAAGTGAAGACTTCATA 7020
DB 7301 GAAACCCGAGCGCGCCCAAGAGAGGCGCGGAGCGGTGAGTGAAGAGACATA 7360
QY 7021 GAGAGTGCCTTCAACAGCTGAGCTTAAGTCTTTGGCAGGCCCCCAAGCGCGAT 7080
DB 7361 GAGAGTGCCTTCAACAGCTGAGCTTAAGTCTTTGGCAGGCCCCCAAGCGGTGAG 7420
QY 7081 TCGAGGCTTTTCAACGAGGAGGAGGAGTTCGCGGAGTGCAGAGCGCTCTGATGAG 7140
DB 7421 TCGAGGCTTTTCAACGAGGAGGAGGAGTTCGCGGAGTGCAGAGCGCTCTGATGAG 7480
QY 7141 TTGGCCCTTTTCAACAGGAGTTCATCTTTCAATGCGGCTTCAGAGGAGGAGCTTGA 7200
DB 7481 TTGACTCTCTTCAACAGGAGTTCATCTTTCAATGCGGCTTCAGAGGAGGAGCTTGA 7540
QY 7201 GATTCAGAGCTTGAAGCTTGAAGAGTGAAGCGGCAACCCCGGCGAGGAGGAGTGA 7260
DB 7541 GATTCAGAGCTTGAAGCTTGAAGAGTGAAGCTTCAAGCTCCCTCAGAGGAGGAGGAGTGA 7600

QY 7261 GCTCCGAGCTCGGAGCTCGAGGCTCTGCTTACTTGTCCGAGGAGGAGCAGCTCGTCTG 7320
DB 7601 GCTCCGAGCTCGGAGCTCGAGGCTCTGCTTACTTGTCTCGAGGAGGAGTGAAGCTCGTCTG 7660
QY 7321 TGTGCTCAGATGTCATATCTCTGAGAGCGGAGCTCTAATACTCTTGTAGTCCGAGAG 7380
DB 7661 TGTGCTCAGATGTCATATCTCTGAGAGCGGAGCTCTAATACTCTTGTAGTCCGAGAG 7720
QY 7381 GAGAAGTACCGATTAACCCCTTGAAGCAGTCCCTGTGGGATATCAACAAGGTGAC 7440
DB 7721 GAAAAGTTCGCAATTAACCCCTGAGCACTCCCTGTGGGATATCAACAAGGTGAC 7780
QY 7441 TGTACCAACAAGAGCGGCTCAGTAAGGCTTAAAAAGTAACTTTGATAGATGCA 7500
DB 7781 TGTACATCATCAAGAGCGGCTCAGTAAGGCTTAAAAAGTAACTTTGATAGATGCA 7840
QY 7501 GTGCTCAGCTCTTACAGCTCAGTCTTAAAGAGATTAAGCTACGCGCTCAAGTCT 7560
DB 7841 GTGCTCAGCGCATTATGACTCAGTATTAAAGATCATCAAGCTACGCGCTCAAGTCT 7900
QY 7561 ACCGCAAGGCTCTCAGCAGTGAAGAGGCTTGCAGTAAACCCCACTTTCGACAG 7620
DB 7901 AGTCAAGGCTCTCAGCAGTGAAGAGGCGTGCAGCTTACCTCACCCCAATTCGACAAA 7960
QY 7621 TCTAAATATGAGGTTTGGAGCTAAGAGGTCGAGGCTTGGCCGAGAGGCGTTAAC 7680
DB 7961 TCCAGATTAAGGTTTGGAGCTAAGAGGTCGAGGCTTGGCCGAGAGGCGTTAAC 8020
QY 7681 ATCAAGTCCGTTGGAAGAGCTCTGAGAGACTCAGAAACACCAATTCACACAACTT 7740
DB 8021 ATCAAGTCCGTTGGAAGAGCTCTGAGAGACTCAGAAACACCAATTCACACAACTT 8080
QY 7741 ATGGCCAAAAATGAGGTTTCTGCGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 7800
DB 8081 ATGGCCAAAAATGAGGTTTCTGCGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8140
QY 7801 CTATGCTTAACTCTGAGCTCGGAGCTGAGGAGTTCGAGAGAGAGTGCCTTATGACAT 7860
DB 8141 CTATGCTTAACTCTGAGCTCGGAGCTGAGGAGTTCGAGAGAGAGTGCCTTATGATGTC 8200
QY 7861 ACAAAAAAATCTCTGAGGAGGAGTGGGAGCTTCTTATGATTCAGATTCCTCCGCT 7920
DB 8201 ACAAAAAAATCTCTGAGGAGGAGTGGGAGCTTCTTATGATTCAGATTCCTCCGCT 8260
QY 7921 CAGCGGTGAAGTCTTCTTGAAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7980
DB 8261 CAGCGGTGAAGTCTTCTTGAAGAGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8320
QY 7981 TATGATACCGAGTCTTGAAGTCAACGCTCACTGAGAGAGACTCAGAGCTGAGAGTCC 8040
DB 8321 TATGATACCGAGTCTTGAAGTCAACGCTCACTGAGAGAGACTCAGAGCTGAGAGTCC 8380
QY 8041 ATATATCGGAGCTGCTCTTGGCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8100
DB 8381 ATATATCGGAGCTGCTCTTGGCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8440
QY 8101 AGACTTAACTGAGAGGAGGAGGAGTTCAGACAGAGAGGAGGAGGAGGAGGAGGAGGAG 8160
DB 8441 AGACTTAACTGAGAGGAGGAGGAGTTCAGACAGAGAGGAGGAGGAGGAGGAGGAGGAG 8500
QY 8161 TGCAGGCGAGCGGAGGAGTTCAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8220
DB 8501 TGCAGGCGAGCGGAGGAGTTCAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8560
QY 8221 GCTTACGAGCTTGAAGCTGAGAGGATTAAGCGCCCAAAATGCTGTATGCGGAGT 8280
DB 8561 GCTTACGAGCTTGAAGCTGAGAGGATTAAGTTCGCCCAAAATGCTGTATGCGGAGT 8620
QY 8281 GACTTGTGTATCTCAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8340
DB 8621 GACTTGTGTATCTCAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8680
QY 8341 TTCACGAGGCTTATGAGCAGGATATTCGCCCCCTCTGTGAGACCCCGGAGAGTAT 8400

Db	8681	TTCCACGAGGCTATAGACCAAGTATTTCTGCCCTTCCGGACACCCCCAGACCGGAATAT	8740
QY	8401	GATCTGAGAGCTGATPAACATCTTTCCTCTCAAAATGTGTGTGGCGCTGGGCCCAAGGC	8460
Db	8741	GACCTGGAGGCTGATPAACATCTTTCCTCTCAAAATGTGTGTGGCGACCGGACCCACAGGCG	8800
QY	8461	CGCGCGAGATCTACTACCTGACAGAGACCCCTTCAACTCCATCGGCCCGGGCTGGGAA	8520
Db	8801	CGCGCGAGATCTACTACTGACAGAGACCCCTTCAACTCCATCGGCCCGGGCTGGGAA	8860
QY	8521	ACAGTTAGACACTCCCTGTTCATTCATGCGTGGAAAATCATTCATCACTACGCCCGAAC	8580
Db	8861	ACAGTTAGACACTCCCTGTTCATTCATGCGTGGAAAATCATTCATCACTACGCCCGAAC	8920
QY	8581	ATATGGGGGTGGCATAGTCTGTATGACACACTTCTCTTCATCTTCATATGGCTCAAGACAG	8640
Db	8921	ATATGGGGGTGGCATAGTCTGTATGACACACTTCTCTTCATCTTCATATGGCTCAAGACAG	8980
QY	8641	CTGGACACGAAACCTCAACTTTGAGATGTACGAGAGGTGTACTCCGTAGTCCCTTGAC	8700
Db	8981	CTGGACACGAAACCTCAACTTTGAGATGTACGAGAGGTGTACTCCGTAGTCCCTTGAC	9040
QY	8701	CTCCACAGTATATTTGAAAAGGTATCATGGGCTTTGACGTTTTTCTCTGCAACATACACT	8760
Db	9041	CTCCACAGTATATTTGAAAAGGTATCATGGGCTTTGACGTTTTTCTCTGCAACATACACT	9100
QY	8761	CCCCACGAACTGACACGGGTGGCTTACAGCCCTTCAGAAAACCTTGGGGGCCACCCCTGAG	8820
Db	9101	CCCCACGAACTGACACGGGTGGCTTACAGCCCTTCAGAAAACCTTGGGGGCCACCCCTGAG	9160
QY	8821	GCGTGGAAAGAGCCCGGACAGCTGACATCAGGCGCTCCATCTCCCGTGGGGGAGAGCG	8880
Db	9161	GCGTGGAAAGAGCTCGGGCGCGTGCAGATCAGGCGCTCCATCTCCCGTGGGGAGAGCG	9220
QY	8881	GCGTTTTGGCGGTGATATCTCTTCAATTTGGGGGGGGAAGACCAAGCTCAACTCACTCA	8940
Db	9221	GCGTTTTGGCGGTGATATCTCTTCAATTTGGGGGGGGAAGACCAAGCTCAACTCACTCA	9280
QY	8941	TTGCGGAGAGGCGCGCTCTTGATTTTATCCAGCTGTTCACCGTGGCGCGCGGGGCG	9000
Db	9281	TTGCGGAGAGGCGCGCTCTTGATTTTATCCAGCTGTTCACCGTGGCGCGCGGGGCG	9340
QY	9001	GACATTTATATCAGCGTGTCCCGTGGCCGAGACCCCGCTTATGTCTTTGGGCTACTCTTA	9060
Db	9341	GACATTTATATCTGTGGTGTCTGCGTCCGAGACCCCGCTTATGTCTTTAGGCTACTCTTA	9400
QY	9061	CTTTTGTAGGGGTAGGCTTTTCTTACTATCCCGCGTCG	9098
Db	9401	CTTTCCGTGGGAGCTAGGCTTTTCTTACTATCCCGCGTCG	9438
RESULT 12			
LOCUS	AF169005	9700 bp	RNA linear
DEFINITION	Hepatitis C virus isolate NDM5 polyprotein gene, complete cds.		
ACCESSION	AF169005		
VERSION	AF169005.1	GI:6707285	
KEYWORDS			
SOURCE	Hepatitis C virus		
ORGANISM	Hepatitis C virus		
REFERENCE	1 (bases 1 to 9700)		
AUTHORS	Kurihara, C., Ishiyama, N., Nishiyama, Y., Fukushima, S., Kageyama, T., Katayama, K. and Miura, S.		
TITLE	Molecular characterization of hepatitis C virus genotype 2a from the entire sequences of four isolates		
JOURNAL	J. Med. Virol. 64 (4), 466-475 (2001)		
MEDLINE	21361470		
PUBMED	11468731		
REFERENCE	2 (bases 1 to 9700)		
AUTHORS	Katayama, K. and Kurihara, C.		

DIRECT Submission
 Submitted (14-JUL-1999) Basic Research Division, Biomedical
 Laboratories, Inc., 1361-1 Matoba, Kawagoe, Saitama 350-1101, Japan
 JRES
 source
 Location/Qualifiers
 1. 9700
 /organism="Hepatitis C virus"
 /mol_type="genomic RNA"
 /isolate="NDM59"
 /db_xref="taxon:11103"
 /note="genotype: 2a"
 341. 9442
 CDS

341_9439.rge

Page 522

TITLE
JOURNAL
FEATURES
Source

Direct Submission
Submitted (14-Jul-1999) Basic Research Division, Biomedical
Laboratories, Inc., 1361-1 Matoba, Kawagoe, Saitama 350-1101, Japan
Location/Qualifiers

1..9700
/organism="Hepatitis C virus"
/mol_type="Genomic RNA"
/isolate="NDM59"
/db_xref="taxon:11103"
/note="genotype: 2a"
341..9442
/codon_start=1
/product="polyprotein"
/protein_id="AAP25613.1"
/db_xref="GI:6707286"
/translation="MSTNTPDKRKTNRTRRPPDVFYVGGGQIVGGVLLPRGRPL
GVRAATKTSERSRGRGRROPIDPKDRSTGSKMCKPYPWLYINEGIGMAQLIPSG
SRSPGPTDPRKSRNVGKVIDITGCFADLMKGIIPVGAPLGVARALAHGVLEDS
GVNVAAGNLPGCSFSIFLLALSCITVPASAVGNKISDYMVTNDSNDSITWQLQA
AVLHVGCVPCEKMGNISRCMIPIVSPNVAVQRGALTOGRRAHIDMWMSATLCBALY
VGLDGGWMLIAQMFIVSPDHMFVQENCSIRYGAITGHRAMDMWMSPTATWLL
AYAMRBEVILIDISGAHMGVMEGLAFPSMGQAKVYVLLTAGVDATRRSIAGSV
AAHSTGLAGLFTSGAKONIQLIINTGSMHNRNALNDSITGLASLPTYRPNSS
GCEBLSACRGIOAFRIQMGSTLAEEDVYTPEDMRPYCCHYPRKQCSIYASRSCEGV
YCEPERSVAVGTTDRLGVPITYTGENEDTVFLINSTRPEGDSFGCTWMSGTFTKTC
GAEPCHIRADFNASMDLCEPTDCEFRKHDPATYIKCSGPMLTRCLVDYPRLMHPC
TINYITFKIRMYGVGEHRLLTAACNFTRGDPCMLIEDRDSQLPLHSTTEMALPCSS
YSPLDPLSTGLHINQIVDOVMGSLPALTYIIRWENVVLLITLADARVCAKMLD
MLLQOAPALAEKLVIVHAASAQNGCFPLVPAAYAVIGRANVPAAYSLTGILM
PCLLLALPOAYAYIDASVHGCTGVALLILITFLITPTEYKILSRSLMWLCYLLTL
GEMIDIEWPMPMGARGRDIITAAATMFCGGVFDITKMLMAGVALLRDLATLRPA
YFRAHALRLCTMVRHLAGRVYOMALALIGMTGYIYIDHLSPMDWASQGLRLDA
VAEPIIFSPMEKVVIVMGCAETACGDVLGEVSRALGREIILGPDGYSKCMRLD
APTAAVQOTRGLGIAIVSMTGRDKTEQAGEIIOVSTYQSFLGTSISGVLVYVGG
AAGKTLASRGPTOMYSASAEGLNOMSPSPGKLSLEPTCGAVDLYLVTRNAVIPA
RRRGDRGALNSRPLSTLKSSGCGPLCPRGAAVCFRAASRGVAKSIITPIVERA
LDVNTSPFESDSTPPAVQIYOVGYLAHPTSGSKTRVPAVAAQGYTLVLPSPV
AALTGAVYLSKAHGINPNIRTVGRVTTGSESITYSYGFLADGGCAGAYDIIICDD
ECHAVDATLILGAVTLDOAETAGVRLTVLATITPQSVTTPPHNIEVALGCGEITP
VYCARPLASTYIKGGRHLIFCHSKKCCDELSALRSMGLNAAVYRGDVAEYIPGQGV
VVYATMALMTGYGQDSDVSDCNVAUTOVDVLPDPTITGTOVPDVAISRQDRGR
TGRGIGITRYVSTGERASGMFDSVULCECYDGAANYELTPSTTYRLAAVPTPGL
PVCDDHLFEWEAVFTGLTHIDAHFLSOTKSGHFALVAYQATVCAKARALSTPVLN
MMCLRLKLEPWEAVFTGLTRIGAPVNEVTLHPVTVYIICQADLEWTSVPMAGV
GVYAAVAAVCLATGCVSIIGRHLNORAVAAPREKYLEFDEMEBCASAAALIECCQ
RIEMKSKIOGLLOOASKOADIOPAVOASVMPREOFYAKHMNFISGLOYLGSITL
VSGNIVASMMAFSAALTSPLSTSTILLINIGSMILNOIAPRGALGVFVSGVAA
VSEIGIGKVLVDILAGYAGISGALVAFKTMSEKBSMEVILNPILBSPLGALVNV
ICAAIIIRHVGBERGAOVMMNRLLIAPSRNHAAPHYVIESDSNQVTQLDGLTIT
SLIRLHMNIITEDCPIPCAGSMLEWMDVCTILTFKMWLTSKLPFKMGLPIISCO
KGRGVWAGTGIWTTTRPCPGANISGNVRLSMIITGKTCMNTWQGFPCINCYBEGC
VPEAPNFKTALIRVVAASEYAEVTOHGSYVIGLTVDLS.KVCOGLSPSPFSVNDGCO
QIRPAPTPKPPFRDEYSPCVGLNSFVSGOICDDPDDADVLTSM.LTDSHITAEAA
ARLARGSPPSBASASSASOLASASLARTCTTHKARALYANVDANLPNGCVTIRESS
KAVVLDLPMABERSDLEPSISETHLPMNRPPALPANARPDNPPVBSKRPDY
QPTPVGSCALPEPKTPTPPRRRRRTVGLNENITGDLQOLIAKAFQPLSGSGSGLIS
TGADADSGSRTPPDSALSETGISISMPELEGPGGPDLEPBEVLEQPPGGEVVP
GSGDSMSTCEEDSDVSCSMYSWMTGALITCSPEEKPIIPNLSNIRYKNKYV
CTTSKASLPAKKTDFDRMVLIVADSVOLITKIASAKSARLLTEBAKOITPPHS
ARSKYFGAKEVATSLSGRNPITTIWAKYKDLLEDSQPIPTTIWAKNVEPCVDPAKGGR
KAARLVYDPUGVRCBKMAVYVTOQLPAVWQASGYFQISQRYEFLIKAMAEKKK
DMPGFSYDTRCFDSVTERDIRTEESLYRCSLPEEHTAIIHSTELVYGGPLNS
GQTCGRVCRASGVLTYSMGTITCYKIALAAKAAIVAPITVYVACDDLVVIESQNS
TEEDENLRATFETAMTRVSPADGPPEPEVLEITSCSNVSAVXAKPOORRYVLL
DPTTLISRAMETVRSVNSMGTINIOVAPTLVWRNVMTWPSIIMOADTLOONLN
FENYGVYISYSPDLPAITIRLGLDAFLSHITYSPHETLVASALRLGADPLPAWKS
RAVAVASLISRGRAAICGRVFNMAVAKTLKLTPIEARLLDLSWFTVAGAGGDTI
YHVSARAPLLLSLTLVGLFLTPAR"

ORIGIN

Query Match	86.6%	Score 7878.6;	DB 14;	Length 9700;
Best Local Similarity	91.6%	Pred. No. 0;		
Matches 8334;	Conservative 3;	Mismatches 762;	Indels 0;	Gaps 0;

QY 1 ATGAGCAAAATCTTAAACCTCAAGAAAAACAAAAGAACCAACCGTGGCCCA 60
| | | | |
Db 341 ATAGACAAATCTTAACTTCAAGAAAAACAAAAGAACCAACCGTGGCCCA 400
| | | | |
QY 61 GACGTTAAGTTTCGGGCGGCGGCGAGATGTTGGCGAGTATATCTTGTCCGCGAG 120
| | | | |
Db 401 GACGTTAAGTTTCGGGCGGCGGCGAGATGTTGGCGAGTATATCTTGTCCGCGAG 460
| | | | |
QY 121 GGGCCAGAGTTGGGTGTGCGCGGACAGAGAACTTCCGAGCGGTCCAGCCAGTGA 180
| | | | |
Db 461 GGGCCAGAGTTGGGTGTGCGCGGACAGAGAACTTCCGAGCGGTCCAGCCAGTGA 520
| | | | |
QY 181 AGGCGCAGGCCATCCCTTAAGATCGGCGCTCCAGTGGCAATCTCGGGAAAAACAGA 240
| | | | |
Db 521 AGAGCCAGGCCATCCCTTAAGATCGGCGCTCCAGTGGCAATCTCGGGAAAAACAGA 580
| | | | |
QY 241 TACCCCTGGCCCTTATACGGGAAATGAGGAGCTCGGCTGGGAGAGATGCTCTGCCCC 300
| | | | |
Db 581 TACCCCTGGCCCTTATAGGGAATGAGGAGCTCGGCTGGGAGAGATGCTCTGCCCC 640
| | | | |
QY 301 CGAGGTTCCCGTCTCTTGGGCCCCCAATGATCCCGGAGTATAGTTCGCAACGTGGT 360
| | | | |
Db 641 CGAGGTTCCCGTCTCTTGGGCCCCCAATGATCCCGGAGTATAGTTCGCAACGTGGT 700
| | | | |
QY 361 AAGGTATGATACCTTAAGATCGGCGCTTGGCGACCTCAATGGGGTATACCTCTGCTG 420
| | | | |
Db 701 AAGGTATGATACCTTAAGATCGGCGCTTGGCGACCTCAATGGGGTATACCTCTGCTG 760
| | | | |
QY 421 GGGCCCCCGCTCGGCGGCGCTCGGAGAGCTCTGCGGATGGCGTGAAGTCTGAGAGAC 480
| | | | |
Db 761 GGGCCCCCGCTTGGTGGCGCTCGGAGAGCTCTGCGGAGCGGTGAAGTCTGAGAGAC 820
| | | | |
QY 481 GGGGTTAATTTGCAACAGGAACTTACCGGTTGCTCTTTTCTATCTTCTTGCGCC 540
| | | | |
Db 821 GGGGTTAATTTGCAACAGGAACTTACCGGTTGCTCTTTTCTATCTTCTTGCGCC 880
| | | | |
QY 541 CTGCTTCTGATCAACACCCCGGCTCGGCTGGCGGAGTGAAGAACTCATACCGGC 600
| | | | |
Db 881 TTAATCTCTGATCAACACCCCGGCTCGGCTGGCGGAGTGAAGAACTCATACCGGC 940
| | | | |
QY 601 TACATGATGATCAACAGCTGCAACATGACAGATTAATCTGGAGCTTCAGGCTGCTG 660
| | | | |
Db 941 TAAATGATGATCAACAGCTGCTTAATGACAGATTAATCTGGAGCTTCAGGCTGCTG 1000
| | | | |
QY 661 CTCACGTCCTCGGAGTCTCGGCTGGAGAAATGGGGAATGATCTCATGCTGAGTA 720
| | | | |
Db 1001 CTCACGTCCTCGGAGTCTCGGCTGGAGAAATGGGGAATGATCTCATGCTGAGTA 1060
| | | | |
QY 721 CCGGTTCAACGGAATGTCGCGGCTGACAGCGCCCGGCTTCACAGAGGCTTGGCGAG 780
| | | | |
Db 1061 CCGGTTCAACGGAATGTCGCGGCTGACAGCGCCCGGCTTCACAGAGGCTTGGCGAG 1120
| | | | |
QY 781 CACATGCAATGTTGATGATGTCGCGAGCTGCTGCTGCTGCTTACGTTAGGGGAGCTG 840
| | | | |
Db 1121 CACATGCAATGTTGATGATGTCGCGAGCTGCTGCTGCTGCTTACGTTAGGGGAGCTG 1180
| | | | |
QY 841 TGGGTTGGGTTGATGCTCGAGCCCAATGTTTATGTTCTGCGCGAGCACTGGTTT 900
| | | | |
Db 1181 TGGGTTGGGTTGATGCTCGAGCCCAATGTTTATGTTCTGCGCGAGCACTGGTTT 1240
| | | | |
QY 901 GTCCAGAGCTGCAATGTTGCTCACTTACCTGTTACCTCACTGAGACACCGGATGGATGG 960
| | | | |
Db 1241 GTCCAGAGAGTCAATGTTGCTCACTTACCTGTTACCTCACTGAGACACCGGATGGATGG 1300
| | | | |
QY 961 GACATGATGATGATGATGTCGCGAGCTACCATGATCTTGGGCTGAGGATGCGTGT 1020
| | | | |
Db 1301 GACATGATGATGATGATGTCGCGAGCTACCATGATCTTGGGCTGAGGATGCGGTT 1360
| | | | |
QY 1021 CCGGAGTCACTTATGACATCACTTACGGGGCTCATTTGGGGCTCATTTGGCTTGGCC 1080
| | | | |
Db 1361 CCGGAGTCACTTATGACATCACTTACGGGGCTCATTTGGGGCTCATTTGGCTTGGCC 1420
| | | | |

QY 1081 TACTTCTATGAGGAGGAGTGGGAAAGTGTGTATCACTTCTGTTGGCGCGG 1140
| | | | |
Db 1421 TACTTCTATGAGGAGGAGTGGGAAAGTGTGTATCACTTCTGTTGGCGCGG 1480
| | | | |
QY 1141 GTGAGCGCGGCAACCATATCTGTTGGGAGTTCGCGCGAGACACCGGCGCTTACC 1200
| | | | |
Db 1481 GTGAGCGCGGATACCGGAGCATTTGGGGTGGGTCGCGGACGCAACAGCGGCTTGGC 1540
| | | | |
QY 1201 AGCTTATTTGACATGGGCCCCCAGGAGAAATTCAGCTGTTTAACCAATGGCAGTGG 1260
| | | | |
Db 1541 GGCCTCTTACATTTCTGGGCTTAAGAGAAATTCAGCTGTTTAACCAATGGCAGTGG 1600
| | | | |
QY 1261 CACATCAACGAGCAGCCCTGAACTGCAATGACTCTGTCGACACCGGCTTTATCGGCT 1320
| | | | |
Db 1601 CACATCAACGAGCAGCTGCTGAAATTCGAAATGACTCTTGAACACCGGCTTTATCGGCT 1360
| | | | |
QY 1321 CTGTTTACACCCACAGCTTCACTGATGAGATGTCGGAACGATGTCGCTGCGC 1380
| | | | |
Db 1661 CTGTTTACACCCACAGCTTCACTGATGAGATGTCGGAACGATGTCGCTGCGC 1440
| | | | |
QY 1381 AGTATGAGGCTTTCGGGTTGGGATGGGCGCTTGCATATATGAGATTAATGACCAAT 1440
| | | | |
Db 1721 GGTATCAGGCTTTCGGGTTGGGATGGGCGCTTGCATATATGAGATTAATGACCAAT 1480
| | | | |
QY 1441 CCAGAGATATAGACCTTATGCTGGGCACTAACCAACAGGAGTGGCGGTCTCC 1500
| | | | |
Db 1781 CCAGAGATATAGACCTTATGCTGGGCACTAACCAACAGGAGTGGCATATGCTCC 1840
| | | | |
QY 1501 GCGAAGACTGTGTGTGGCCAGTGTACTGTTTCAACCCGAGCCAGTGTAGTGGGACG 1560
| | | | |
Db 1841 GCGAAGCTGTGTGTGGCCAGTGTACTGTTTCAACCCGAGCCAGTGTAGTGGGACG 1900
| | | | |
QY 1561 ACCGACAGCTTGGAGCGCCCACTTACAGTGGGGGAGATATAGACAGATGCTTCTTA 1620
| | | | |
Db 1901 ACCGACAGCTTGGAGCTGCTACTTACAGTGGGGGAGATATAGACAGATGCTTCTTA 1960
| | | | |
QY 1621 TTGAACAGACTGACACACCGCTGGGGTATGGTGGCGACGAGTGAATCTGTT 1680
| | | | |
Db 1961 TTGAACAGACTGACACACCGGGGAGTATGGTGGCGACGAGTGAATCTGTT 2020
| | | | |
QY 1681 GGCTACACCAAGACTTGGGCGGACACACCTGCGTACTTAAAGACTGATCAACGCGAGC 1740
| | | | |
Db 2021 GGTTTACCAAGACTTGGGCGGACACCGGCTGCGATTAAGACTGATCAATGCGAGC 2080
| | | | |
QY 1741 ACGGACCTGTTGTGCCCCAGACTGTTTAAAGACATCTGATACCACTTACCTCAAA 1800
| | | | |
Db 2081 ACGGACCTGTTGTGCCCCAGACTGTTTAAAGACATCTGATACCACTTACCTCAAA 2140
| | | | |
QY 1801 TGGGCTCTGGGCGCTGGGCTCAACGCAAGTGTGATGATCAACCTTACAGGCTTGG 1860
| | | | |
Db 2141 TGGGCTCTGGGCGCTGGGCTCAACGCAAGTGTGATGATCAACCTTACAGGCTTGG 2200
| | | | |
QY 1861 CATTAACCTGACAGTAACTATACATCTTCAAAATTAAGATGATGTTGGAGGGGTT 1920
| | | | |
Db 2201 CATTAACCTGACAGTAACTATACATCTTCAAAATTAAGATGATGTTGGAGGGGTT 2260
| | | | |
QY 1921 GAGCAGAGGCTCAACGCTGATGCAATTTCACTGTTGGGAGTGTGTTGCACTTGAAGAG 1980
| | | | |
Db 2261 GAGCAGAGGCTCAACGCTGATGCAATTTCACTGTTGGGAGTGTGTTGCACTTGAAGAG 2040
| | | | |
QY 1981 AGAGACAGAGTCAATGTCCTGTTGTTGCACTTCAACAGGAGTGGGCACTTTAACT 2040
| | | | |
Db 2321 AGAGACAGAGTCAATGTCCTGTTGTTGCACTTCAACAGGAGTGGGCACTTTAACT 2080
| | | | |
QY 2041 TGTCTTAACTCGGACCTGCGGCTTGTGACTGTCTTCTCAACCTTCAACCAATC 2100
| | | | |
Db 2381 TGTCTTAACTCGGACCTGCGGCTTGTGACTGTCTTCTCAACCTTCAACCAATC 2140
| | | | |
QY 2101 GTGAGAGTCAATTTCAATGATGCTTATGACCTGCTTCAACCAAAATGATGCTGAGTGG 2160
| | | | |
Db 2441 GTGAGAGTCAATTTCAATGATGCTTATGACCTGCTTCAACCAAAATGATGCTGAGTGG 2200
| | | | |
QY 2161 GAGTGGTAACTCTTATCTGCTTATGAGGAGCGCAGGAGTTGGCGCTTATGAG 2220
| | | | |

Db 2501 GAGGAGGTAGTACTTATTCCTGCTTCGCGATCCAGGGGTCTGCGCTTCCTGTGG 2560
Qy 2221 ATGCTCATCTTGTGGGCGAGGCGAAGACACTAGAGAAGTGTGATCTTGGACGCT 2280
Db 2561 ATGCTCATCTTGTGGGCGAGGCGAAGACACTAGAGAAGTGTGATCTTGGACGCT 2620
Qy 2281 GCGAGCGAGCTAGCTGCAATGAGCTTCTATATTTGTCACTTTTCTGTGGCTGTGG 2340
Db 2621 GCGAGCGAGCTAGCTGCAATGAGCTTCTATATTTGTCACTTTTCTGTGGCTGTGG 2680
Qy 2341 TACATCAAGGTGCGAGTATGCTTCTTACCTATTCCTTCACTGCTGTGTCTTT 2400
Db 2681 TACATCAAGGTGCGAGTATGCTTCTTACCTATTCCTTCACTGCTGTGTCTTT 2740
Qy 2401 AGCTTACTGCTCTAGGATGCGCCCAAGGCTTATGACGATCTGAGCATGCG 2460
Db 2741 TGCTGTGCTCTCTAGGATGCGCCCAAGGCTTATGACATGATGATGCTGACGGA 2800
Qy 2461 CAGATAGAGCGGCTCTGTGTATGATGATCTCTTTACTCTCACTCCCGGGTATAG 2520
Db 2801 CAGATAGAGCGGCTCTGTGTATGATGATCTCTTTACTCTCACTCCCGGGTATAG 2860
Qy 2521 ACCCTTCTCAGCGGCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
Db 2861 ATCTCTCTCAGCGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2920
Qy 2581 GTCCAGAGTGGGACCAACCATGACAGGTGCGCGGTGAGTGGCATCATAGGGCC 2640
Db 2921 ATCCAGAGTGGGACCAACCATGACAGGTGCGCGGTGAGTGGCATCATAGGGCC 2980
Qy 2641 GTCCCATATTTCTACCCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2700
Db 2981 GCCACATGTTCTGCGCGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3040
Qy 2701 GGGCTCTTACCTCTCTTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2760
Db 3041 GGGCTCTTACCTCTCTTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3100
Qy 2761 GCTTACTGAGATGTGACCATGAGCAAGCATCTCGCGGGGAGAGTACGTCCAGATG 2820
Db 3101 GCTTCTGTGAGTGTGTGACATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3160
Qy 2821 GCGCTACTAGCCCTTGGCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2880
Db 3161 GCGCTACTAGCCCTTGGCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3220
Qy 2881 TCGGATTTGGGCTGT 2940
Db 3221 TCGGATTTGGGCTGT 3280
Qy 2941 AGTCCGATGAGAGAAAGTCAATTGTCTGGGGAGCGGAGCAGTGTCTGTGGGGAATT 3000
Db 3281 AGTCCGATGAGAGAAAGTCAATTGTCTGGGGAGCGGAGCAGTGTCTGTGGGGAATT 3340
Qy 3001 TTACACGAGACTTCCGCTGTCCGCCGACTTGTGTGGAGGTCTCTTGTGGCCGACTGAT 3060
Db 3341 TTGACACGAGACTTCCGCTGTCCGCCGACTTGTGTGGAGGTCTCTTGTGGCCGACTGAT 3400
Qy 3061 GGTCTATCTTCCAGAGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3120
Db 3401 GGTCTATCTTCCAGAGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3460
Qy 3121 CGTGGCCCTTTTGGGACCAATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3180
Db 3461 CGAGGTCTCTTGGGCGCTATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3520
Qy 3181 GGGGAAATTCAGGTCTGT 3240
Db 3521 GGGGAAATTCAGGTCTGT 3580
Qy 3241 GTTTTGTGACTGTCTATCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3300

Db 3581 GTCTTATGTGACTGTTTATCATGTAGAGCTGTGCAATAAACTTAGCCGCTGCGGGCCCG 3640
Qy 3301 GTCCAGCAGATGATCTCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3360
Db 3641 GTCCAGCAGATGATCTCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3700
Qy 3361 ACTAAATCTTTGAGCCGT 3420
Db 3701 ACCAAATCTTTGAGCCGT 3760
Qy 3421 GCTGATGTATCCCGCTGT 3480
Db 3761 GCTGATGTATCCCGCTGT 3820
Qy 3481 CCTCTTCCACTGTGAAGGGGTCTCTCAAGAGCCCGGTCTATGCCCCAGGGGCGACGT 3540
Db 3821 CCTCTTCCACTGTGAAGGGGTCTCTCAAGAGCCCGGTCTATGCCCCAGGGGCGACGT 3880
Qy 3541 GTCCAGTCTTCCGGGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3600
Db 3881 GTCCAGTCTTCCGGGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3940
Qy 3601 CCCGTGTGACACTGTGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3660
Db 3941 CCCGTGTGACACTGTGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4000
Qy 3661 CCTGTGTGCTCCAGCATATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3720
Db 4001 CCGGCGGTGCTCAGACCTATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4060
Qy 3721 AGCACCAAGTCTGT 3780
Db 4061 AGTACCAAGTCTGT 4120
Qy 3781 TCAGTGTGCTCCAGCCCTGT 3840
Db 4121 TCAGTGTGCTCCAGCCCTGT 4180
Qy 3841 AACATTAGAGCTGT 3900
Db 4181 AACATTAGAGCTGT 4240
Qy 3901 GGGCAATTCCTGT 3960
Db 4241 GGGCAATTCCTGT 4300
Qy 3961 GAATGTGATCCGT 4020
Db 4301 GAATGTGATCCGT 4360
Qy 4021 GAGACAGCTGT 4080
Db 4361 GAGACAGCTGT 4420
Qy 4081 ACCCCCCACCCCAATAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4140
Db 4421 ACCCCCCACCCCAATAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4480
Qy 4141 GGGAGGGGCAATTCCTGT 4200
Db 4481 GGGAGGGGCAATTCCTGT 4540
Qy 4201 AAGAAAGT 4260
Db 4541 AAGAAAGT 4600
Qy 4261 TACTACAGAGGT 4320
Db 4601 TACTACAGAGGT 4660
Qy 4321 ACCGAGCCCTCATGTGACGGGTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4380
Db 4661 ACCGAGCCCTCATGTGACGGGTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4720

4381 GGGGTCACTCAAGTTGAGACTTCAGTTAGACCCCACTTCAACATTAACACAGATT 4440
4721 GGGGTCACTCAAGTTGAGACTTCAGTTAGACCCCACTTCAACATTAACACAGACT 4780
4441 GTCCCTCAAGACCTGTCTCAGCTAGACGCGCGGGGTGCGACGAGGTAGAGGAAGT 4500
4781 GTCCCGCAAGACCTGTCTCAGCTAGACGCGCGGGGTGCGACGAGGTAGAGGAAGT 4840
4501 GGCATTATATGATATGTTTCACTGTGAGAGCTCAGGAATGTTTGAAGTATAGT 4560
4841 GGCATTATATGATATGTTTCACTGTGAGAGCTCAGGAATGTTTGAAGTATAGT 4900
4561 CTCTGTGAGTGTACAGACGAGGGGCGCGATGATATGAGTCAACCATGAGAGACACC 4620
4901 CTCTGTGAGTGTACAGACGAGGGGCGCTTGTGATGAGTCAACCATGAGAGACTAC 4960
4621 GTGAGGCTCAGGCGGTATTTCAACACGCGCTTGTGCTGTGCGAAGACATCTTGAG 4680
4961 GTGAGGCTTATGAGCGTATTTCAACACGCGCTTGTGCTGTGCGAAGATCATCTTGAG 5020
4681 TTTTGGAGGCGAGTTTTCACCGGCTTCACACATAGATGCCCATCTTCTTCCAAACA 4740
5021 TTTTGGAGGCGAGTCTTTACTGTGCTTCACACATAGATGCCCATCTTCTTCCAAACA 5080
4741 AAGCAATCGGGGGGAAAAATTTGCGATTAATTAACAGCCCTACAGGCTACAGTGTGCGTAG 4800
5081 AAGCAATCGGGGGGAAAAATTTGCGATTAATTAACAGCCCTACAGGCTACAGTGTGCGTAG 5140
4801 GCGAAAGCCCCCGCGCTCTGAGGAGTATGAGTGAAGTGTGACTGCACTCAAGCCC 4860
5141 GCGAAAGCCCCCGCGCTCTGAGGAGTATGAGTGAAGTGTGACTGCACTCAAGCCC 5200
4861 ACACTGTGGGCCCCACACCTCTCTGTACCGCTTGGGCTGTGTACCAACGAGTCAAC 4920
5201 ACACTGTGGGCCCCACACCTCTCTGTACCGCTTGGGCTGTGTACCAACGAGTCACT 5260
4921 CTGACACATCGCGTGAAGGAATATCATGCGACCTGCACTGCAAGCGGACCTTGAAGTATG 4980
5261 CTGACACATCGCGTGAAGGAATATCATGCGACCTGCACTGCAAGCGGACCTTGAAGTATG 5320
4981 ACCAGACATGAGGTCTTGTGAGGAGGAGTCTTGGCGGCGCTGCGCGGTATTTGCTGGCG 5040
5321 ACCAGACATGAGGTCTTGTGAGGAGGAGTCTTGGCGGCGCTGCGCGGTATTTGCTGGCG 5380
5041 ACCGAGTGTGTTTGCATCATCGGCGGCTTGCACATTAACACGAGCGCGTGTGCGCG 5100
5381 ACCGAGTGTGTTTGCATCATCGGCGGCTTGCACATTAACACGAGCGCGTGTGCGCG 5440
5101 GACAAAGAGTCTCTTATGAGGCTTTGATGAGATGGAAGATGTGCTTATGAGCGGCT 5160
5441 GACAAAGAGTCTCTTATGAGGCGTTGACGAGATGGAAGATGTGCTTATGAGCGGCT 5500
5161 CTGATTGAAAGAGGAGCGCGGATAGCGAGATGCTGAAGTCAAGATCCAAGGCTTATG 5220
5501 CTGATTGAAAGAGGAGCGCGGATAGCTGAAGTCAAGATCCAAGGCTTATG 5560
5221 CAGCAAGCTTCCAAACAGCTCAAGATCAACCCACTGTGAGGCTTCAATGCGCCAG 5280
5561 CAGCAAGCTTCCAAACAGGCGCGGATATCAACCCCGTGTGAGGCTTCAATGCGCCAG 5620
5281 GTGAGAACATTTGAGGCGCAACATGATGAACTTCAATTAAGCGGATCCAATATCTGCA 5340
5621 GTGAGAACATTTGAGGCGCAACATGATGAACTTCAATTAAGCGGATCCAATATCTGCA 5680
5341 GGAATATCAACACTGCGCAGGGAACCTCTGAGTATGCTTCAATGAGGCTTCAATGAGC 5400
5681 GGAATATCAACACTGCGCAGGGAACCTCTGAGTATGCTTCAATGAGGCTTCAATGAGC 5740
5401 CTGACCAAGTCCGCTGTCAACAGACCACTATCTTCTCAACATTTTGGGGGCTGTGCTA 5460
5741 CTGACCAAGTCCGCTGTGTGAGTCAACCACTATCTTCTCAACATTTTGGGGGCTGTGCTG 5800

5461 GCATCCCAATTGCAACACCGCGGGGGCCACTGCGTGTGCTGATGAGGCA 5520
5801 GCGTCCCAATTGCG 5860
5521 GCTGCGTATGAGCAGTATAGGCTTATGAGTATGAGTATGAGTATGAGTATGAGTATG 5580
5861 GCTGCGTATGAGCAGTATAGGCTTATGAGTATGAGTATGAGTATGAGTATGAGTATG 5920
5581 GCGGCGATTTGCGGGGCTCTGCTGCGATTCAGATCATATGCTGTGCGAAGGCGCTCATG 5640
5921 GCGGCGATTTGCGGGGCGCGCTGAGATTCAGATCATATGCTGTGCGAAGGCGCTCATG 5980
5641 GAGATGTGCTCAACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5700
5981 GAGATGTGCTCAACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6040
5701 ATCTGCGCGCGCACTTCTGCGCGCGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCG 5760
6041 ATCTGCGCGCGCACTTCTGCGCGCGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCG 6100
5761 AATGACTATTTGCTTGTGCTTCAAGAGAAATCAAGTGTGCGCGCGCGCGCGCGCGCG 5820
6101 AATGACTATTTGCTTGTGCTTCAAGAGAAATCAAGTGTGCGCGCGCGCGCGCGCGCG 6160
5821 GAGTGTGAGTGTGCGAGCGTGTGAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5880
6161 GAGTGTGAGTGTGCGAGCGCGTGTGAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6220
5881 CTGAGAAAGTCTCAACATGATTAATGAGAGCTGCGCGCGCGCGCGCGCGCGCGCGCG 5940
6221 CTGAGAAAGTCTCAACATGATTAATGAGAGCTGCGCGCGCGCGCGCGCGCGCGCGCG 6280
5941 CTGCGGAGTGTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 6000
6281 CTGCGGAGTGTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 6340
6001 TCCAAATTAATTTCCCAAGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6060
6341 TCCAAATTAATTTCCCAAGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6400
6061 GGGGTGTGGCGCGCATGAGCATGATGACCAACGAGTGTCTTGTGCGCGCGCGCGCG 6120
6401 GGGGTGTGGCGCGCATGAGCATGATGACCAACGAGTGTCTTGTGCGCGCGCGCGCGCG 6460
6121 GGCATATGTGCGCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 6180
6461 GGCATATGTGCGCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 6520
6181 CAGGGGACCTTCTCTATGATTTGTAACGAGAGGCGAGTGTGCGCGCGCGCGCGCG 6240
6521 CAGGGGACCTTCTCTATGATTTGTAACGAGAGGCGAGTGTGCGCGCGCGCGCGCGCG 6580
6241 AACTTAAAGTGTGCGCATGAGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6300
6581 AACTTAAAGTGTGCGCATGAGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6640
6301 GGGTCAATCACTATCACTAAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 6360
6641 GGGTCAATCACTATCACTAAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 6700
6361 CCTCTGCGGAGTCTTCTTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 6420
6701 CCTCTGCGGAGTCTTCTTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 6760
6421 CCGAAGCGTTTTTCCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 6480
6761 CCGAAGCGTTTTTCCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 6820
6481 GGGTCCAGTGTCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 6540
6821 GGGTCCAGTGTCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 6880
6541 GATCACTTCTATCACTGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 6600

Db 6881 GATCCGTCATATACAGCGCGAGGCTGACGCGCGCTTAGAGCGGGATCAACCCCG 6940
Qy 6601 TCCGAGGCAAGCTCTCGGCGAGCGCAGCTATCGGCACCATCGCTGGAGGACCTCGACC 6660
Db 6941 TCCAGAGAGGCTCTCGGCGAGCGCAGCTATCGGCACCATCGCTGGAGGACCTCGACC 7000
Qy 6661 ACCCAGCGCAAGCCTATGATGTGAGACATGGTGGATCTTAACCTTTTCATGGGCGCGAT 6720
Db 7001 ACCCAGCGCAAGCCTATGATGTGAGACATGGTGGATCTTAACCTTTTCATGGGCGCGAT 7060
Qy 6721 GTGACTGGATAGAGTCTGGGCTCCAAAGTGTGTCTTGGACTCTTCGACCCCAATGTC 6780
Db 7061 GTGACTGGATAGAGTCTGGGCTCCAAAGTGTGTCTTGGACTCTTCGACCCCAATGTC 7120
Qy 6781 GAAGAAAGAGAGCAGCTTGGAGCTTCGATACATCAGAAATACATGCTCCCAAGAGAG 6840
Db 7121 GAAGAAAGAGAGCAGCTTGGAGCTTCGATACATCAGAAATACATGCTCCCAAGAGAG 7180
Qy 6841 TTCCACACGCTTACCAGGCTTGGAGCAGCGCTGATTTACAACCCACCGCTTGTGAATCG 6900
Db 7181 TTCCACACGCTTACCAGGCTTGGAGCAGCGCTGATTTACAACCCACCGCTTGTGAATCG 7240
Qy 6901 TGGAAAAGGCGAGATTACCAACCGGCACTGTTCGGGCTGTGCTTCCTCTCTAGG 6960
Db 7241 TGGAAAAGGCGAGATTACCAACCGGCTGTGCTTCCTCTCTAGG 7300
Qy 6961 AAAACCCGAGCGCTCCCGCCCAAGAGAGGCGCGACAGTGGGCTTAAGTGAAGATCCATA 7020
Db 7301 AAGACTCGAGCGCTCCCGCCCAAGAGAGGCGCGACAGTGGGCTTAAGTGAAGATCCATA 7360
Qy 7021 GGAGATGCGCTTCAACAGCTGGCCATTAACTCTTTGGCCAGCCCGCCCAAGCGCGAT 7080
Db 7361 GGAGATGCGCTTCAACAGCTGGCCATTAACTCTTTGGCCAGCCCGCCCAAGCGCGAT 7420
Qy 7081 TCAGGCTTTTCAACGCGGCGCGGCGCTGCGCATTCGCGCAGTCAAGAGCGCTCTGATAG 7140
Db 7421 TCAGGCTTTTCAACGCGGCGCGGCGCTGCGCATTCGCGCAGTCAAGAGCGCGCTCTGATAG 7480
Qy 7141 TTGCGCCCTTGGAGAGCAGGTTTCATCTCTTCATGCGCGCGCGCTTGAAGGAGAGCTTGA 7200
Db 7481 TCGGCTCTTTGGAGAGCAGGTTTCATCTCTTCATGCGCGCGCGCTTGAAGGAGAGCGCGG 7540
Qy 7201 GATCAGAGCTGAGGCTGAGCAGGTAGAGCCCAACCCCGCCCAAGCGGCGGTGCA 7260
Db 7541 GATCAGAGCTGAGGCTGAGCAGGTAGAGCCCAACCCCGCCCAAGCGGCGGTGCA 7600
Qy 7261 GCTCCCGGCTGGAGCTCGGGGTCTGTGTCTACTTGTCTCCAGAGAGAGATCGCTGTGTG 7320
Db 7601 GCTCCCGGCTGGAGCTCGGGGTCTGTGTCTACTTGTCTCCAGAGAGAGATCGCTGTGTG 7660
Qy 7321 TGCTGCTCATGTATATCTCTGAGACGCGGCTCTTAATTAATCTCTTGTAGTCCGAGAG 7380
Db 7661 TGCTGCTCATGTATATCTCTGAGACGCGGCTCTTAATTAATCTCTTGTAGTCCGAGAG 7720
Qy 7381 GAGAAGTTACCGATTAAACCCCTTGAACAACCTCCCTTGGATATCAACAAGATGTAC 7440
Db 7721 GAAAAGTTGCAATTAACCCCTTGGATATCGGCTATGCGATTCACAAATTAAGTGTAC 7780
Qy 7441 TGTACACAAACAAAGAGGCTCTACTAAGGCTTAAAGGTTAACTTTTGTAGATGCA 7500
Db 7781 TGTACTACATCAAAAGAGGCTCTACTAAGGCTTAAAGGTTAACTTTTGTAGATGCA 7840
Qy 7501 GTGCTCGACTCTACTAGACATCAGTCTTAAGAGCATTAAAGTACAGCGGCTCCCAAGTTC 7560
Db 7841 GTGCTCGAGCGCTCTTAATTAATGACTCTTAATTAATGAGCTTAAGCGGCTCCCAAGTTC 7900
Qy 7561 ACCGAAAGGCTCTCACCATTGAGAGAGGCTTGCAGTTAAACCCACCCCATTTGCAAGA 7620
Db 7901 AGCGAAAGGCTCTCACCATTGAGAGAGGCTTGCAGTTAAACCCACCCCATTTGCAAGA 7960
Qy 7621 TCTAATATGAGGTTTGGGGCTTAAGAGGTCGCAAGCTTGTCCGGAGGGCGCTTAAACAC 7680

Db 7961 TCCAAATATGAGTTTGGGGCTAAGAGAGTCCGACGTTGTCCGGAGGGCGCTCAACAC 8020
Qy 7681 ATCAATCCGTTGAGAGAGACCTCTCGAGAGACTCGAAGACCAATTTCCCAACCAT 7740
Db 8021 ATCAATCCGTTGAGAGAGACCTCTCGAGAGACTCGAAGACCAATTTCCCAACCAT 8080
Qy 7741 ATGACCAAAATGAGGTGTTCTGCTGAGACCCCAACCAAGGGGCGCAAGAGAGCTCGC 7800
Db 8081 ATGACCAAAATGAGGTGTTCTGCTGAGACCCCAACCAAGGGGCGCAAGAGAGCTCGC 8140
Qy 7801 CTTATGTTTACCTGAGACCTCGGCGTCAAGGCTGCGAGAGATGAGCCCTTATGACAT 7860
Db 8141 CTTATGTTTACCTGAGACCTCGGCGTCAAGGCTGCGAGAGATGAGCCCTTATGAGAT 8200
Qy 7861 ACACAAAACCTCTCAGCGGCTGATGAGGCGCTTCTTATGATTCAGATTTCCCGCT 7920
Db 8201 ACACAAAACCTCTCAGCGGCTGATGAGGCGCTTCTTATGATTCAGATTTCCCGCT 8260
Qy 7921 CAGCGGTAGAGTTTCTCTTGAAGCATGCGGAGAAAGAGAGACCTTATGGGTTTTTG 7980
Db 8261 CAGCGGTAGAGTTTCTCTTGAAGCATGCGGAGAAAGAGAGACCTTATGGGTTTTTG 8320
Qy 7981 TATGATACCCGATGCTTGAATCAACCGTCACTGAGAGAGACATCAGTACCTGAGAGTCC 8040
Db 8321 TATGATACCCGATGCTTGAATCAACCGTCACTGAGAGAGACATCAGTACCTGAGAGTCC 8380
Qy 8041 ATATATCGGCGCTGCTCTTGGCCGAGAGAGCCACACTGCGATACACTCGTAACTGAG 8100
Db 8381 ATATATCGGCGCTGCTCTTGGCCGAGAGAGCCACACTGCGATACACTCGTAACTGAG 8440
Qy 8101 AGACTTACGTTGAGAGGCTTATGTTCAACAGCAAGGCTCAACCTGCGGCTTCAAGCGT 8160
Db 8441 AGACTTACGTTGAGAGGCTTATGTTCAACAGCAAGGCTTCAACCTGCGGCTTCAAGCGT 8500
Qy 8161 TGCAGGCGAGCGGGGTGCTCAACCACTAGATGGGGACACCATCATGACTACGTGAA 8220
Db 8501 TGCAGGCGAGCGGGGTGCTCAACCACTAGATGGGGACACCATCATGACTACGTGAA 8560
Qy 8221 GCTTATGCGGCTTGAAGCTGAGGATTAATGCGCCCAACATGCTGTATGCGCGAT 8280
Db 8561 GCTTATGCGGCTTGAAGCTGAGGATTAATGCGCCCAACATGCTGTATGCGCGAT 8620
Qy 8281 GACTTGTGTCTTCAAGAAAGCCAGGAGACCGAGAGAGAGAGAGGAACTGAGAGCC 8340
Db 8621 GACTTGTGTCTTCAAGAAAGCCAGGAGACCGAGAGAGAGAGAGGAACTGAGAGCC 8680
Qy 8341 TTACGAGAGGCTATGACAGAGTATTCGCGCCCTCGTGTGACCCCGCCAGACGAGAT 8400
Db 8681 TTACGAGAGGCTATGACAGAGTATTCGCGCCCTCGTGTGACCCCGCCAGACGAGAT 8740
Qy 8401 GATCTGAGGCTATTAATCTTGTCTCTCAATGTGTCTGTGCGCTGAGCCCAAGAGC 8460
Db 8741 GATCTGAGGCTATTAATCTTGTCTCTCAATGTGTCTGTGCGCTGAGCCCAAGAGC 8800
Qy 8461 GCGCGAGATCTACTGACAGAGACCTTCAATCTCAATGCGCGGCTGCTTGGAA 8520
Db 8801 GCGCGAGATCTACTGACAGAGACCTTCAATCTCAATGCGCGGCTGCTTGGAA 8860
Qy 8521 ACAGTTAGACCTCCCTGCTGAAATTCATGCGTGGGAAACATCATCAGTACGCGCGGAC 8580
Db 8861 ACAGTTAGACCTCCCTGCTGAAATTCATGCGTGGGAAACATCATCAGTACGCGCGGAC 8920
Qy 8581 ATATGAGGCTCGCATGCTCTGATGACACACTTCTCTCATTTCTCAATGCTCAAGACG 8640
Db 8921 ATATGAGGCTCGCATGCTCTGATGACACACTTCTCTCATTTCTCAATGCTCAAGACG 8980
Qy 8641 CTGAGACAGACCTCAACTTTGAGATGAGAGCGGTGTACTCTCGTGAATCCCTTGAC 8700
Db 8981 CTGAGACAGACCTCAACTTTGAGATGAGAGCGGTGTACTCTCGTGAATCCCTTGAC 9040
Qy 8701 CTCGAGCTTATTTGAAGGTTTACATGAGGCTTGAAGGCTTTTCTCTGACACATCACT 8760
Db 9041 CTCGAGCTTATTTGAAGGTTTACATGAGGCTTGAAGGCTTTTCTCTGACACATCACT 9100

QY 8761 CCCACGAAGTCAACAGGGGCTTACGCCCCCTGAGAAAGCTGGGGGCGCCACCCCTAG 8820
 Db 9101 CCCACGAAGTCAACAGGGGCTTACGCCCCCTGAGAAAGCTGGGGGCGCCACCCCTAG 9160
 QY 8821 GCGTGAAGAGCCGGGACGTCAGTCAGGAGCGTCCCTCATCTCCCGTGGGGAGAGCG 8880
 Db 9161 GCGTGAAGAGTCGGGGCGCTGAGTAAAGGCGTCCCTCATCTCCCGTGGAGGAGAGCG 9220
 QY 8881 GCGGTTTGGGCTGATATCTCTTCAATTGGGCGGTGAAGACCAAGCTCAAACTCACTCCA 8940
 Db 9221 GCGATTGGCGGTGGATATCTCTTCAACTGGGCGGTGAAGACCAAGCTCAAACTCACTCCA 9280
 QY 8941 TTGCGGGAAGCGGCGCTCTGATATTTATCCAGCTGGTTCACGTCGGGCGCGGGGAGG 9000
 Db 9281 TTGCGGGAAGCGGCGCTCTGATATTTATCCAGCTGGTTCACGTCGGGCGCGGGGAGG 9340
 QY 9001 GACATTATCAGACGCTGTGCGGTCGCCGACCCCGCTTATGCTTTTGGCTTACTCTTA 9060
 Db 9341 GACATTATCAGACGCTGTGCGGTCGCCGACCCCGCTTATGCTTTTGGCTTACTCTTA 9400
 QY 9061 CTTTGTGTAGGGTGAAGGCTTTTCTTCTACTCCCGGCTGG 9099
 Db 9401 CTTTGTGTAGGGTGAAGGCTTTTCTTCTACTCCCGGCTGG 9439

RESULT 13

AF238482 9416 bp RNA linear VRL 26-MAR-2000
 LOCUS Hepatitis C virus 2a polyprotein gene, complete cds.
 AF238482
 VERSION AF238482.1 GI:7329202

SOURCE

Hepatitis C virus
 Hepatitis C virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE 1 (bases 1 to 9416)
 Itakura,J., Nagayama,K., Enomoto,N., Kurosaki,M., Watanabe,H. and Sato,C.

TITLE Full length cDNA sequence of HCV genotype 2a, strain MD2a-2
 JOURNAL Unpublished
 AUTHORS Itakura,J., Nagayama,K., Enomoto,N., Kurosaki,M., Watanabe,H. and Sato,C.

REFERENCE 2 (bases 1 to 9416)
 Submitted (24-FEB-2000) Second Department of Internal Medicine, Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-Ku, Tokyo 113-8519, Japan

TITLE Direct Submission
 JOURNAL
 AUTHORS Sato,C.

FEATURES
 source location/Qualifiers
 1..9416
 /organism="Hepatitis C virus"
 /mol_type="genomic RNA"
 /strain="MD2a-2"
 /db_xref="taxon:11103"
 /note="genotype: 2a"
 307..9408
 /codon_start=1
 /product="polyprotein"
 /protein_id="AF59941.1"
 /db_xref="GI:7329203"
 /translation="MSTNPKDKTKRNTNRPRDVKFPGGQIVGGVILVPRGRPLGVATRTSBRSGPRRODIPKDRSTGSMKPCPYPMPLVYENBGLGNAWMLISRLSRPSMPTPRRKRNVGKVIDLTGCFADLMKGTIPVGAFLGVAALAHGVRVLEDGNAVATGNLPGCSFSIFLLALLSCVTPVSAVQVKNSTSETWINDCSNDSITWOLQAVLHAGVCVPCERTSRMIPVSPVAVQVQPGALTOGLRTIHDVWVMSALCSALYVGLCGGVLAQLFIYVSPRHFVQECNSIYPGATIGHRMAMMMMSPTTIVLVAVMARVBEVIDIISGAGVWFMGLAFPSOGAMAVAVILLLTAGVBARHTTGSVA GTTSGPAGIFTSQPKRONIOLINTNSMHNPTALCONSLNGLNGPMAALPYTNRPSGCBERSKXKNIAAFRIKMETLQEDVYTPNEDRPYCHYPRKQCGIFPAGSVCPVYCTPSVAVGTGNNKLVPTPTWGENEDVDFILNSTRPPGVSFGCTMNSSTFTYTCGAPRCIRADFNASTDLCEPTDCFRKHPETATYKCSGSPWLTRCLVDYPRLMHYPC TVNYSIFKIRMYGVGVHRLTAACNFSRGRCNLEDRDSQLTRPLHSTTEMAILPSCS

CDS

1..9416
 /organism="Hepatitis C virus"
 /mol_type="genomic RNA"
 /strain="MD2a-2"
 /db_xref="taxon:11103"
 /note="genotype: 2a"
 307..9408
 /codon_start=1
 /product="polyprotein"
 /protein_id="AF59941.1"
 /db_xref="GI:7329203"
 /translation="MSTNPKDKTKRNTNRPRDVKFPGGQIVGGVILVPRGRPLGVATRTSBRSGPRRODIPKDRSTGSMKPCPYPMPLVYENBGLGNAWMLISRLSRPSMPTPRRKRNVGKVIDLTGCFADLMKGTIPVGAFLGVAALAHGVRVLEDGNAVATGNLPGCSFSIFLLALLSCVTPVSAVQVKNSTSETWINDCSNDSITWOLQAVLHAGVCVPCERTSRMIPVSPVAVQVQPGALTOGLRTIHDVWVMSALCSALYVGLCGGVLAQLFIYVSPRHFVQECNSIYPGATIGHRMAMMMMSPTTIVLVAVMARVBEVIDIISGAGVWFMGLAFPSOGAMAVAVILLLTAGVBARHTTGSVA GTTSGPAGIFTSQPKRONIOLINTNSMHNPTALCONSLNGLNGPMAALPYTNRPSGCBERSKXKNIAAFRIKMETLQEDVYTPNEDRPYCHYPRKQCGIFPAGSVCPVYCTPSVAVGTGNNKLVPTPTWGENEDVDFILNSTRPPGVSFGCTMNSSTFTYTCGAPRCIRADFNASTDLCEPTDCFRKHPETATYKCSGSPWLTRCLVDYPRLMHYPC TVNYSIFKIRMYGVGVHRLTAACNFSRGRCNLEDRDSQLTRPLHSTTEMAILPSCS

ORIGIN

Query Match 86.5%; Score 7869.8; DB 14; Length 9416;
 Best Local Similarity 91.5%; Pred. No. 0;
 Matches 8330; Conservative 0; Mismatches 769; Indels 0; Gaps 0;
 QY 1 ATGAGCAAAATCTTAACCTCAAAAGAAAACCAAAAGAAACCAACCGTCCGCCACAA 60
 Db 307 ATGAGCAAAATCTTAACCTCAAAAGAAAACCAAAAGAAATCTTAACCGTCCGCCACAG 366
 QY 61 GACGTTAAGTTCCGGGCGGGCGGCGACGATCTGTTGGGAGATATCTTTGTTCCGGCGAG 120
 Db 367 GACGTTAAGTTCCGGGCGGGCGGCGACGATCTGTTGGGAGATATCTTTGTTCCGGCGAG 426
 QY 121 GGGCCAGAGTTGGGCTGGCGCGGACAAAGSAAAGCTTGGAGGGGTCCGACGACGCGGA 180
 Db 427 GGGCCAGAGTTGGGCTGGCGCGGACAAAGSAAAGCTTGGAGGGGTCCGACGACGCGGG 486
 QY 181 AGCGCGACGCCATCCCTAAAGATCGGCGCTCCACTGGCAAAATCTGGGGGAAAACAGGA 240
 Db 487 AGCGCGACGCCATCCCTAAAGATCGGCGCTCCACTGGCAAAATCTGGGGGAAAACAGGA 546
 QY 241 TACCCCTGGCCCTATACGGAAATGAGGAGCTGGCTGGGACAGATGGCTCTGTCCTCCC 300
 Db 547 TACCCCTGGCCCTATATAGGAAATGAGGAGCTGGCTGGGACAGATGGCTCTGTCCTCCC 606
 QY 301 CGAGGTTCCCGGCTCTTGGGGGCGCCCAATGACCCCGGATAGGTGGCGCAACGTCGGCT 360
 Db 607 CGAGGTTCCCGGCTCTTGGGGGCGCCCAATGACCCCGGATAGGTGGCGCAACGTCGGCT 666
 QY 361 AAGGTATGATACCTTAACGTCGGCTTGGCGGACCTGATGGGGGTATCATCTCTGTCGTG 420

YSDLPALSTGLHONIVDVYVWGLTPALTKVYVWEMVVLFIILADARVACIM
 MLLIIOABALEKIVVILHAASAASNGLLYPIFPVAAATYIGRAVPLAAYLITGLM
 PFCULILVPEQAYAYDASVHGVGAALLILFPLTPEYKILSRCLMWLLCYLTL
 GENLIQEWPEPTQARGRDGIIMAAITFCGVGVFDITKMLAVLAPYALIRDLTRP
 YVFAHALLMTCVNLNLAGKYVOMALAGVMTSTYIYDHLTPMSDMAASGLRDLA
 VAVEPIIFSMKEKVIAMGAETACGDIKHGLPVSRLGEILILPGADGYSAGRLIA
 APITAAOORTGLGALVYVMTGRDTEOAGELOVSTVOSFTGTTISGLVMTVYHG
 AGNKTLAGRGPRVYOMYSAGEIDLKMGSPRGKSLIEPCGCAVDLYLTRNADVIPA
 RRRGDRCALISPRPLSTLKSGSGGYLCEPRGAIVCI FRAVYCSRGVAKSIDPI PBT
 LDIIVTSPTFSDNSTPFAVPOYQVETILHATSGSGSTKTPVAVYAGGVVLINBVS
 AATLFGAALSKAHGPNPRTGVRVTVTLAGAATYSTYGEFLADGGCAGAVYIILICD
 ECHADVSTIILIGVLDQAEITGVLVATVATPGSVYTPPNILEVILAGGGEIP
 PYGAPILSHIKGRHLIFCHSKKCDIELAALRGGLANAVAYRGLAEVILPQGV
 VVPAATMLTGTGDFDSVIDCNAVATVODVADLDEPTITOTVTPDADSRGRRBR
 TGRGRGIYRYVSTGRASGMPDSVVLCEYVDGAAWYELTPAETRYLAAPRTPL
 PVCODHLEFMEAVFTGLTHIDAFISQITQSGSNFAYLVAVQAVCARAPAPPSMV
 MMCLTRLEKPLTGPPLVRLGSVANEITLHPVYKLAITCMQADLEVMTSTWVLG
 GVLAVAAYACLATGCVSIIIGRLHINORAVAPDKEVLYEAFDEMEBCASRALVEBQ
 RIAMELSKSIOGLQOASKOADIPQAVOASMPKVEOPAKHMMNFISGLOYLAGLST
 LPGNPAVASMMAFSAALTSPLSTTILINLIIGKMLASOIAAPAGATGFGVSGLVGA
 VGSIGKXVLVDILAGYAGISGALVAFKIMSEKSMEDVYVNLBGLISPGALVYGV
 ICATILRRHYGREGAVQWNNRLIAPASGNHVAIPHVYTESDASQVTOLOLSLIT
 SLILRLHNMITBEDCPIPCAGSWLIRDVMDVCTLLDPKMWLTSKLPPKMGDLPEIISQ
 KGKGYWAGTGTMTTRPCPGANISGVNVLGSMKLTGPKTCMTWOGTFPIYNCYBGC
 VPKAPNFKVIAIWRVASEVYAEVTOHGSYSYITGLTDMKVPCCOLSPFFFSWVDCV
 QIHRFAPTPKPPFRDEVSFCVGLNSFVSGSLPCDPEPTDVLVLSMTLDTSTIETAT
 ARLLAGSPSEBSASSASQASASLBACTTHKTYTDVMDVADS.LMGDDVYRTIBES
 KVVVLDLDPMAEESDLBESIPSEYMLSNRPPLPAARADVPNPLVESKRPDY
 QPATAVAGCALPEPKKAPTPRRRRRTVGLSESTIGALQOALIKSGOEPPOGSEVLS
 TGDADADSGSRSPDELALETGTSMSPLBEPDPLDEPOVPOPGCEVAVP
 GSDSGMSTCSSEENVFVFCISISYMTGGLITPCIEPEKELPENPLSNLIDNKKY
 CTTSKASLSRAKVTDFRMQVLDAYVDSVLIKOIKLAASKVLSARLLTLEBACOLTPHS
 ASKTYFGAKEVRSLSGRAVNHKISKWKLDEBSQPIPTTTAKKBEPCVDNPKGK
 KAARLIVPDLGVRCERKALVDYQKLDQAVMGASVGYSPQGVVEFLYKMAEKK
 DPMGFSYDTRCEPSTVERDITRESITYPCSLPEBAHTAHSLETLRVLYGGMFSK
 GQACGYRRCRASGVLTYSMGNTTTCVYKLAACAKAIAVAPMLVCGDDLYVSESG
 TEBDERNLRAFTGAMTRGAPGDPPEPEKDELITGSCSNVSVHIGQORRRVYLTR
 DPTPTIARAAMETVHRSVNSMIGNTIOVAPTYVAMVMTTFEFTSLMODTDLNIN
 FEMTGSVYSVSLDLPATILBERLGLDAFSLTISPHBLRVAAALRLKCAPRIARAKS
 RAAVAVASLISHGGRVAVCGRYLFNNAVATKLTPLPEARHLIDLSWFTVGAGGHI
 YHSVSRARBLRLGLLFLVGVFLPLPAR"

Dh 667 AAGTCATCGATACCCCTAACGTGTGCTTTCGCACTCTAGGGATGACATCCCCGTGTA 726
Qy 421 GCGGCCCGCGCTCGGCGGCGCTCGCCAGAGCTCTGCGCATGGCGGTGAGAGCTCGAGAGAC 480
Dh 727 GGGGCCCACTTGGTGGCGCTTGCAGAGCTCTGCGCGACGGCGTGAAGTCTCGAGAGAC 786
Qy 481 GGGGTATATTTTGGCAAGGAACTTACCCGGTTGCTCTTTTCTATCTTTCTTGGGCC 540
Dh 787 GGGGTATATTTGCAACAGGAACTTACTGTTGTCTCTTTTCTATTTTCTTGGGCC 846
Qy 541 CTGCTGCTCGCATACACCCCGGTCTCGGTGCGCAAGTGAAGAACATCACTCGGC 600
Dh 847 CTACTGCTCGCGCACCGTTCGGTCTGTGTCTCAAGTGAAGAACACCACTGAGAAC 906
Qy 601 TACATGGTGACTAACGACTGCACCAATGACAGCATTACTGAGAGCTCAAGGCTGTGC 660
Dh 907 TACATGGTGACTAATGACTGTTCCAAATGACAGCATCACTGAGAGCTTACGGCGGCTC 966
Qy 661 CTCACGCTCCCGGGTGCCTCCGTCGAGAAAGTGGGAAATGCATCTCAGTGCTGATA 720
Dh 967 CTCACGCTCCCGGGTGCCTCCGTCGAGAAAGTGGGAAATGCATCAAGTGTGATA 1026
Qy 721 CCGGTCTCACGAAATGAGCGCTGACAGCGGCGCGGCGCTCAAGGAGGCTTGGAGAG 780
Dh 1027 CCGGTTCACCAACGAGCGCTGACAGAGCGCGCGCTCAAGGAGGCTTGGAGAG 1086
Qy 781 CACATCGACATGGTGTGATGTCGCGCACGCTGTCTGCTCCCTCTAAGTGGGAGACTC 840
Dh 1087 CACATCGACATGGTGTGATGTCGCGCACGCTGTCTGCTCCCTCTAAGTGGGAGACTC 1146
Qy 841 TCGGCTGGGTGATGCTCGAGCGCAATGTTCAATGTCTTCGCGGACGACCATGTTT 900
Dh 1147 TCGGCGCGGGTGGTGTGTCGAGCGCAAGCTGTTCATGTCTCGCGCGCCCTCATGTTT 1206
Qy 901 GTCGAAGACTGCAATTTGCTCATCTACCTGATACCACTGACACCGGAGTGGAGAG 960
Dh 1207 GTGAGGAATGCAATTTGCTCATCTATCTGTGTGACATCACTGACACCGTATGGAGAG 1266
Qy 961 GACATGATGATGACTGTGCGCCACGAGCTACCATGATCTTGGGCTGAGGAGTGCCTG 1020
Dh 1267 GACATGATGATGACTGTGCGCCACGAGCTACCATGATCTTGGGCTGAGGAGTGCCTG 1326
Qy 1021 CCGGAGTCAATTAAGACATTAATGCGGGGCTCATTTGGGGCTGATGTTTGGCTTGGCC 1080
Dh 1327 CCGGAGTCAATTAAGACATTAATGCGGGGCTCATTTGGGGCTGATGTTTGGCTTGGCC 1386
Qy 1081 TACTTCTATGACGGAGGAGGTGGGCGAAGTGTGTCAATCTTCTGTTGCGCGCGGG 1140
Dh 1387 TACTTCTATGACGGAGGAGGTGGGCGAAGTGTGTCAATCTTCTGTTGCGCGCGGG 1446
Qy 1141 GTGAGCGCGGCAACCACTACTGTTGGGGGTTCTGCGCGGAGACCAACCGGCGCTCAC 1200
Dh 1447 GTGAGCGCGGCAACCACTACTGTTGGGGGTTCTGCGCGGAGACCAACCGGCGCTCAC 1506
Qy 1201 AGCTTATTTGACATGGGCGCGGAGCAAAATCCAGCTCGTTAAACCAATGGAGCTGG 1260
Dh 1507 GGCATCTTCAACCTCCGCGCTTAAGCAGAACATTCAGCTCATTAACCAATGGAGCTGG 1566
Qy 1261 CACATCAACGCGACCGGCGTGAATGCAATGACTCTTGGACACCGGCTTATCGGCT 1320
Dh 1567 CACATCAACGCGACCGGCGTGAATGCAATGACTCTTGGACACCGGCTTATCGGCGC 1626
Qy 1321 CTGTTTCAACCAAGCTTCACTCGTACAGATGTCCGAAACGATGTCGCTGCGC 1380
Dh 1627 TTGTTTCAACCAAGAACTTCACTCGTACAGATGTCCGAAACGCTGTCCGCTGCGC 1686
Qy 1381 AGTATCGAGGCTTCCGGGTGGAGTGGGGCGCTTGGCAATGAGAAATATGTCACCAAT 1440
Dh 1687 AACATCGAAGCTTCCGGATAGAGTGGGGCACTTGGAGTACGAGAGTATGTTAACCAAT 1746
Qy 1441 CCAAGAGATATGAGACCTATGCTGGAGTACCAACCAAGGAGGAGTGGGAGTGTCC 1500
Dh 1747 CCAAGAGATATGAGACCTATGCTGGAGTACCAACCAAGGAGGAGTGGGAGTGTCC 1806

Qy 1501 GCGAAGACTGTGTGTGGGCCCAAGTGTACTGTTTCAACCCCGGCGCAAGTGTAGTGGAGAG 1560
Dh 1807 GCGGGGTGTGTGTGGGCCCAAGTGTACTGTTTCAACCCCGGCGCAAGTGTAGTGGAGAG 1866
Qy 1561 ACCGACAGGCTTGGAGCGCCCACTTAACAGTGGGGGAGATGAGACAGATGTTCTTCTA 1620
Dh 1867 ACCAACAACCTTGGAGTGGCCCACTTAACAGTGGGGGAGATGAGACAGATGTTCTTCTA 1926
Qy 1621 TTGAACAGACTGCACACCGGCTGGGTATGGTTGCGGTGACCGTGAATGAATCTTCT 1680
Dh 1927 TTGAACAGACTGCACACCGGCGGGGTATGGTTGCGGTGACCGTGAATGAATCTTCACT 1986
Qy 1681 GGTACACCAAGACTTGGCGGCGCACACCGCTGCTGACTAGAGACTTCAACGCGCAGC 1740
Dh 1987 GGTTCACCAAGACTTGGCGGCGCACACCGCTGCTGACTAGAGACTTCAATGCGCAGC 2046
Qy 1741 ACGGACCTGTTGTCCCGCACGAGCTGTTTGAAGAGACTGATACCACTTAACCTCAAA 1800
Dh 2047 ACGGACCTTGTGTCCCGCACGAGCTGTTTGAAGAGACTGATACCACTTAACCTCAAA 2106
Qy 1801 TGGGCTCTGGGCGCTGGGCTGACGCGCAAGGAGTGTGATGACTACCGCTTACAGGCTGTG 1860
Dh 2107 TGGGCTCTGGGCGCTGGGCTGACGCGCAAGGAGTGTGATGACTACCGCTTACAGGCTGTG 2166
Qy 1861 CATTAACCCCTGACAGTAACTAATACATCTTCAAAATGAAGATGTATGTGGAGGGGTT 1920
Dh 2167 CATTAACCCCTGACAGTAACTAATACATCTTCAAAATGAAGATGTATGTGGAGGGGTT 2226
Qy 1921 GAGCAGAGGCTCAAGGCTGATGCAATTTTCACTGTTGGGAGTCTTGGCACTTGAAGAGAC 1980
Dh 2227 GAGCAGAGGCTCAAGGCGCGGAGCAATTTCTCTGGGGAGATCTGTGCAACTTGAAGAGAC 2286
Qy 1981 AGAGACAGAAAGTCACTGTCTCTTGTGCACTCCACCAAGGAAATGGGCAATTTAAGT 2040
Dh 2287 AGAGACAGAAAGTCACTGTCTCTTGTGCACTCCACCAAGGAAATGGGCAATTTAAGT 2346
Qy 2041 TGTCTTACTCGGACCTGCGCGGCTTGTGCACTGTGTTCTGCACTCCACCAAAATC 2100
Dh 2347 TGTCTTACTCGGACCTGCGCGGCTTGTGCACTGTGTTCTGCACTCCACCAAAATC 2406
Qy 2101 GTGGAAGTACATTAATGATGAGGCTTATGACCTGCGCTCACAAAATATCTGCGAGATG 2160
Dh 2407 GTGGAAGTACATTAATGATGAGGCTTATGACCTGCGCTCACAAAATATGCTGCGAGATG 2466
Qy 2161 GAGTGGGTAACTCTTATTTCTGCTCTTACGCGAGCGCAGGAGTGTGCGCTGATAGG 2220
Dh 2467 GAGTGGGTAACTCTTATTTCTGCTCTTACGCGAGCGCAGGAGTGTGCGCTGATAGG 2526
Qy 2221 ATGCTCATCTTGTGGGCAAGGCGAAGAGCACTAGAGAACTGTGCACTTTGACGCT 2280
Dh 2527 ATGCTCATCTTGTGGGCAAGGCGTGAAGAGCACTAGAGAAAGTGTGCTTTGACGCT 2586
Qy 2281 GCGAGCGAGCTAGCTGCAATGAGCTTCTATATTTTGTGATCTTTTGTGGGCTGTGG 2340
Dh 2587 GCGAGCGAGCTAGCTGCAATGAGCTTCTGATATTTTGTGATCTTTTGTGGGCTGTGG 2646
Qy 2341 TACATCAAGGAGTGGGTAGTCCCTTAGCTACCTATTTCCCTCACTGAGCTGTGCTT 2400
Dh 2647 TACATCAAGGAGTGGGTAGTCCCTTAGCTACCTATTTCCCTCACTGAGCTGTGCTT 2706
Qy 2401 AGCTTACTGCTCTAGCAATTTGCCCAACAGGCTTATGCTTATGACGATCTGTGATGGC 2460
Dh 2707 TGCCTACTACTCTTACTACTGCGCCCAACAGGCTTATGCTTATGATGATCTGTCAGCGA 2766
Qy 2461 CAGATAGAGCGGCTGCTGCTGATATGATCACTCTTTAATCTTCAACCCCGGGGTAAAG 2520
Dh 2767 CAGATAGAGCGGCTTGTGTTAATATGATTAACCTCTTCAACCTTCAACCCCGGGGTAAAG 2826
Qy 2521 ACCCTTCAAGCGGCTTGTGTTGATGATGATCTTCTGACCTGAGGAGAGGCTATG 2580
Dh 2827 ACCCTTCAAGCGGCTTGTGTTGATGATGATCTTCTGACCTGAGGAGAGGCTTGG 2886

QY	2581	ATCCAGAGATGGGACCAACCTATGCAGGTGGCGGCGTGGCCGTATATGGCATATATGGGCC	2640
Dp	2887	ATCCAAAGATGGGATACCAACCAACGAGCAACGCGTGGCCGCGATATGCATATATGGGCC	2946
QY	2641	GTCCGCATATATCTACCCAGGTGTGGTGTATGACAATACCAAGTGGCTTTGGCGGTCTT	2700
Dp	2947	GCTACATATGTTCTGCCCCGGGTGTAGTGTATTAACATACCAAGTGGCTTTGGCGGTCTT	3006
QY	2701	GGGCGCTGCTTACCTCTTAAAGATGCTTTGACGCGCGTGGCGGTACTTTCGACGGGCTCAC	2760
Dp	3007	GGGCGCTGCTTACCTCTTAAAGATGCTTTGACGCGCGTGGCGGTACTTTCGTCAGAGCTCAC	3066
QY	2761	GCTCTACTGAGAGATGTGCACCATGGCAAGGACATCTCGGGGGGGCAGGTACGTCCAATG	2820
Dp	3067	GCTCTGCTGAGAGATGTGCACCGTGGTGGAACTCGGGGGGGGTAAGTATGTCCAGATG	3126
QY	2821	GGCGCTACTAGCCCTTGGCAGGTGGACCTGGACCTTATCATCTATGACACCTCACCCCTATG	2880
Dp	3127	GGCGCTTATAGCCCTTGGCAGGTGGACCGGACCTTATCATCTATGACACCTCACCTATG	3186
QY	2881	TCGGATTTGGGCTGTAGTGCCCTGCGGGACCTGGCGGTGCGCGTGTGAGCCCTATCATCTTC	2940
Dp	3187	TCGGACGTGGGCTGTAGTGCGGCTGCGGGACCTTGGCGGTGCGGTGTGAGCCCATCATCTTC	3246
QY	2941	AGTCCGATGGAAGAAAGAAATCATTTGTCTGGGAGACGAGACAGCTGTTGTGGGACATT	3000
Dp	3247	AGTCCGATGGAAGAAAGAAATCATGCGCTGGGGAGACAGAGACGCGCATGCGGGGATATC	3306
QY	3001	TTTACACGAGACTTCCCGTGTCCGCCGACTTGGTGGGAGGTCCTTTCGGCCGACGTAT	3060
Dp	3307	TTTACACGAGACTTCCCGTGTCCGCCGACTTGGTGGGAGGTCCTTTCGGCCGACGTAT	3366
QY	3061	GGCTATATCCTTCAAGGGGTGGAGTCTTTCGCCGCCCATCACTAGCTGTACGCCACACAGA	3120
Dp	3367	GGCTATATCCTTCAAGGGGTGGAGGCTTTCGCCGCCCATCACTGTTACGCTACACAGACA	3426
QY	3121	CGTGGCCTTTTGGGACCATATGTGTGAGCATGACCGGGGCGGCAAGACAGAAACAGCT	3180
Dp	3427	CGAGGCGCTTTCGGGCGGTATATGTGTGTGAGCATGACGGGGCGGCAAGACAGAAACAGGCC	3486
QY	3181	GGGGAATTTACAGTCTCTGTCCACAGTCACTCAAGTCTTTCCTCGGGAACATCCATCTCGGGG	3240
Dp	3487	GGGGAATTTCAAGTCTCTGTCCACAGTCACTCAAGTCTTTCCTCGGGAACATCCATTTCCGGG	3546
QY	3241	GTTTGTGTGAGCTGTCTACATGAGAGCTGGCAACAAGACTCTGGCGGCTCACGGGATCCG	3300
Dp	3547	GTTTGTGTGAGCTGTTACACAGAGCTGGCAACAAGACTCTAGCCGGCTCACGGGATCCG	3606
QY	3301	GTCACGAGATGTATCTCCAGTGTCTGAGGGGAGCTTATGATGGGTGGCCACGCCCTCTGGG	3360
Dp	3607	GTCACGAGATGTATCTCGAGTGTCCGAGGGGAGCTTATGATGGGTGGCCACGCCCTCTGGGG	3666
QY	3361	ACTTAATCTTTGGAGCCGTGCACTGTGTGAGCGGTCCGACTGTATCTGTATCTGTGCACGGGAC	3420
Dp	3667	ACCAAACTTTTGGAGCCGTGTATCTGTGTGAGCGGTGTCACTGTATCTGTGTGCACCGGAAC	3726
QY	3421	GCTATGTGCATCCCGGCTCGAGAGACGGGGGACAAAGGGGAGGCTTATCTTCCCGAGA	3480
Dp	3727	GCTATATGTGCATCCCGGCTCGAGAGACGGGGGACAAAGGGGAGGCTTATCTTCCCGAGA	3786
QY	3481	CCTCTTTTCCACTTGAAGGGGTCTTCAGAGAGCCCGGTGTATGCCACAGGGGACACGCT	3540
Dp	3787	CCTCTTTTCCACTTGAAGGGGTCTTCAGAGGGGACCGGGGCTTTGCTCCCTTACGGGGACACGCT	3846
QY	3541	GTCGAGTCTTCCGGGACAGCTGTGTGTCTCTGGGGCGTGTAGTCCATATGATTTTATC	3600
Dp	3847	GTCGAGATTTTCCGGGACCGGTGTGTCTCTGGGGCGTGTAGTCCATATGATTTTATC	3906
QY	3601	CCCGTTGAGACACTGCACATCGTCAAGCGGCTCCGCCACCTTATGTGACAAACAGCACACA	3660
Dp	3907	CCCGTTGAGACACTGCACATCGTCAAGCGGCTCCGCCACCTTCAATGTACAAACAGCACACA	3966
QY	3661	CTGCTGTGCCCAAGACCTATCAGGTGGGATCTTTCATGATGCCGCCGACTGGCAGTGTGAAG	3720

Db	3967	CCACCTGTGCCCCAGACCTTACAGAGTGGGTACTTGCACGCCCCGACTGGCACCGGAAA	4026
OY	3721	AGACACCAAAGTTCCTGTGCGATATATGCTCTCAGGGGTATAAATGTAGTACTTAATCCC	3780
Db	4027	AGTACCAAAGTCCCCGTGCGGTACGCCGCTCAGGGGTACAAAGTGTAGTCTTAATCCC	4086
OY	3781	TCAGTGCCTGCACCTCTGGGGTATTGGGGCGTACTTGTCTAAGGCACATGGCATCAATCCC	3840
Db	4087	TCGGTGCCTGCACCTTGTGGGGTATTGGGGCGTACTTGTCCAAGGCACATGGCATCAATCCC	4146
OY	3841	AAACATTAGAGTGTAGTACAGGACTGTGACACCGGGGGCGCCATACGTACTCCACATAT	3900
Db	4147	AAACATCAGAGATGTGGGTCTAGAGACTGTGACACCGGGGGCGCCATACATCTCCACATAT	4206
OY	3901	GGCAAAATTCCTCGCCGATGGGGGCTGTGCGGGCGCGCCTACACATCATATATGTGAT	3960
Db	4207	GGCAAAATTCCTCGCCGATGGGGGCTGTGCGGGCGCGCCTATGACATCATATATGTGAC	4266
OY	3961	GAATGCCATGCGCGTGACTCTACCAACCATCTTGGCATCGGAAACAGTCTTGTATCAAGA	4020
Db	4267	GAATGCCACGCTGTGATTTCTACCAACCATCTTGGCATTTGGAACAGTCTTGTACCAAGA	4326
OY	4021	GAGCACTGTGGGTCAACCTTACTGTGTGGCTTACAGCTACGGCCCTGGGTCAAGTACA	4080
Db	4327	GAGCACCCGGGTGACGCTTACCTGTATGTCGACACCGCTACGCCCCGGGTGTGTAGCG	4386
OY	4081	ACCCCCAACCCCAATAGAGAGAGTGGCCCTTGTGGGACAGAGGGCGAGATCCCCCTCTAT	4140
Db	4387	ACCCCCAACCCCAATATAGAGAGAGTGGCCCTTGTGGGACAGAGGGGTAGATCCCCCTCTAT	4446
OY	4141	GGGAGGGCGATTCCTCCCTGTCTTACATCAAGAGAGAAAGACATCTGATCTTTCGCATTTCA	4200
Db	4447	GGGAGGGCGATTCCTCCCTGTCTTACATCAAGAGAGAGAACATCTTATTTTCTGCACCTCA	4506
OY	4201	AAAGAAAAGTGTACAGAGCTTCGCGGGCGGCCCTTGTGGGGTATGGGCTTGAACCTAGTGGCA	4260
Db	4507	AAAGAAAAGTGTACAGAGCTTCGCGGGCGGCCCTTGTGGGGACATGGGCTTGAACGCTGTGGCA	4566
OY	4261	TACTACAGAGGGTGTGACGCTCTCCGTATATCAAACTCAGGGAGACGTAGTGTCTGCGCC	4320
Db	4567	TATTACAGAGGGTGTAGACGCTCTCCGTATATCAAACTCAGAGAAATGTAGTGTCTTTCGCC	4626
OY	4321	ACCGAGCCCTCATGACAGGGGTATCTGGGGACTTTGACCTCGGTATCGACTGCAACGTA	4380
Db	4627	ACCGAGCCCTCATGACAGGGGTATCTGGGGACTTTGAGACTTGTGATCGATTGCAACGTA	4686
OY	4381	GCGGTCACTCAAGTTGTAGACTTCAGTTTAGACCCCAATTCAACCATTAACCAACAGATT	4440
Db	4687	GCGGTCACTCAAGTGTAGACTTCAGTTTAGACCCCACTTCAACCATTAACCAACAGACT	4746
OY	4441	GTCCTCTAAGACGCTGTCTACGTAGCAGCGCGCGGGGTTCGACCGGTACGGGAAGACTG	4500
Db	4747	GTCCTCTAAGACGCGCTGTCTGCGTAGCAGCGCGCGGGGTACAGGTATGGGGAAGACTG	4806
OY	4501	GGCAATTATAGTATGTTCACACTGTGTGAGACGAGCCTCAGGAATGTTTGAACAGTGTAGTG	4560
Db	4807	GGCAATTATAGTATGTTCACACTGTGTGAGACGAGCCTCAGGAATGTTTGAACAGTGTAGTG	4666
OY	4561	CTCTGTAGTGTCTACGACGACGAGGGCGCATGTGTATGAGCTCAACCATCGAGACCAACC	4620
Db	4867	CTCTGTAGTGTCTACGACGACGAGGGCGCATGTGTATGAGCTCAACACAGACAGAGACTACC	4926
OY	4621	GTCAGGCTCAGGGGCTTATTCACAGCCCGGGTTCCTGTGTGCAAGACCAATCTTTGAG	4680
Db	4927	GTCAGGCTCAGAGGATTTTCAACACGCGACGGCTGTCTGTGTGCAAGACCAATCTTTGAG	4986
OY	4681	TTTTGGAGGACAGTTTTCACACGGGCTTCACACATAGATGCGCATCTTCTTCCAAACA	4740
Db	4987	TTTTGGAGGACAGTTTTCACACGGGCTTCACACATAGATGCGCATCTTCTTCCAAACA	5046
OY	4741	AAGCAATCGGGGAAAATTTGCAATCTTAAACAGCCTACAGGCTACAGTGTCCGTAGC	4800

Db 5047 AAGCAATCGGGGAGAAATTTGCGATCTTGGTAGCCTATCAAGCCACAGTGTGCTAGA 5106
Qy 4801 GCCAAGCCCCCCCCCGCTCTGGGACGTGATGTGAAGTGTGTACTGACTGAAGCC 4860
Db 5107 GCCAAGCCCCCCCCCGCTCTGGGACGTGATGTGAAGTGTGTACTGACTGAAGCC 5166
Qy 4861 ACACTGTGGGCCCCCAACCTCTCTGTACCGCTTGGGCTGTGTTACCAAGAGTCAAC 4920
Db 5167 ACGCTGTGGGCTCCACACCTCTCTGTACCGTGTGGGCTCGTTCGCAAGATCAAC 5226
Qy 4921 CTCAACATCCCGTGAACGAATATACATCCGCACTGCAAGAGCCGACTTGAAGTCAAG 4980
Db 5227 CTCAACACCTCTGTGAACGAATATACATCCGCACTGCAAGAGCTTGAAGTCAAG 5286
Qy 4981 ACCAGACATGGGTCTTGGCAGAGGAGAGTCTTGGCGGCGCTGCGCCGCTATTCGCG 5040
Db 5287 ACTACACGTGGGTCTTGGCAGAGGAGAGTCTTGGCGGCGCTGCGCCGCTATTCGCG 5346
Qy 5041 ACCGGGTGTGTTGATCATCGGCGCTTGCACATTAAACAGCGAGCGTGTGCGCG 5100
Db 5347 ACCGGGTGTGTTGATCATCGGCGCTTGCACATTAAACAGCGAGCGTGTGCGCG 5406
Qy 5101 GACAGAGAGTCTCTATGAGGCTTTTATGATGAGTGAAGAAATGTGCTTAGAGGCGCT 5160
Db 5407 GACAGAGAGTCTCTATGAGGCTTTTATGATGAGTGAAGAAATGTGCTTAGAGGCGCT 5466
Qy 5161 CTATTTGAAGAGGCGCAGCGATAGCCGAGATGCTGAAGTCCAAAGCTTAAATG 5220
Db 5467 CTCGTGAAGAGGCGCAGCGATAGCCGAGATGCTGAAGTCCAAAGCTTAAATG 5526
Qy 5221 CAGCAAGCTTCCAAACAAAGCTCAGACATACAAACCCGCTGTGAGGCTTCAATGCGCAAG 5280
Db 5527 CAGCAAGCTTCCAAACAAAGCTCAGACATACAAACCCGCTGTGAGGCTTCAATGCGCAAG 5586
Qy 5281 GTAGAAACAATTTGGGCCAAACACATGTGAACTTTAGCGGATCCAAATACCTGCGCA 5340
Db 5587 GTAGAAACAATTTGGGCCAAACACATGTGAACTTTAGCGGATCCAAATACCTGCGCA 5646
Qy 5341 GGAATATCAACACTGCAAGGAAACCTGCAAGTCTTCCATGATGAGGCTTCAAGTCCGCC 5400
Db 5647 GGAATATCAACACTGCAAGGAAACCTGCAAGTCTTCCATGATGAGGCTTCAAGTCCGCC 5706
Qy 5401 CTCACAGTCCGCTGTCAACAAACACATCTCTTCTCAACATTTTGGAGGCTGAGCTA 5460
Db 5707 CTCACAGTCCGCTGTCAACAAACACATCTCTTCTCAACATTTTGGAGGCTGAGCTA 5766
Qy 5461 GCATCCCAATGCAACAACCGCGGGGGCACTGGCTTCGTTTCAGTGGCTTAGTGGGA 5520
Db 5767 GCATCCCAATGCAACAACCGCGGGGGCACTGGCTTCGTTTCAGTGGCTTAGTGGGG 5826
Qy 5521 GCTGCCTAGAGCAGTATAGGCTTAGGTAAAGTGTAGTGAATCTCTGGCAGAGGTATGT 5580
Db 5827 GCAAGCTTAGAGCAGTATAGGCTTAGGTAAAGTGTAGTGAATCTCTGGCAGAGGTATGT 5886
Qy 5581 GCGGGCATTTGCGGGGCTCTGTCGATTCAAAGATCATGTGTGCGAGAAAGCCTTCATG 5640
Db 5887 GCGGGCATTTGCGGGGCTCTGTCGATTCAAAGATCATGTGTGCGAGAAAGCCTTCATG 5946
Qy 5641 GAGATGTGTCAACTGTGCTGGAATTCGTCTCCGGGGCTTGTGTGTGGGAATC 5700
Db 5947 GAGATGTGTCAACTGTGCTGGAATTCGTCTCCGGGGCTTGTGTGTGGGAATC 6006
Qy 5701 ATCTGCGCGGCATTTCTGCGCGCAGACGTGGAGCCGGGGAAAGCGCGCTCCAAATGATG 5760
Db 6007 ATCTGCGCGGCATTTCTGCGCGCAGACGTGGAGCCGGGGAAAGCGCGCTCCAAATGATG 6066
Qy 5761 AATTAAGCTATTTGCTTTGCTTCAAGAGAAATCAGTGTGCGCCCAACCACTAGCTGAG 5820
Db 6067 AATTAAGCTATTTGCTTTGCTTCAAGAGAAATCAGTGTGCGCCCAACCACTAGCTGAG 6126
Qy 5821 GAGTGGAGTGGTGGAGGCTGAGCCCAACTATTTGGCTGCTTACATTAACAGAGCTG 5880
Db 6127 GAGTGGAGTGGTGGAGGCTGAGCCCAACTATTTGGCTGCTTACATTAACAGAGCTA 6186

Qy 5881 CTCAGAAAGCTCCAACTGGATTTACTGAGAGCTGCCCCATCCATGCGGCGCTGTGG 5940
Db 6187 CTCAGAGAGCTCCAACTGGATTTACTGAGAGAGCTGCCCATCCCATGCCCGGTGTGG 6246
Qy 5941 CTCGCGAGTGTGGAGCTGGGTTTGGACCATCTTAACAGACTTTTAAATTTGGCTGACC 6000
Db 6247 CTCGCGAGTGTGGAGTGTGGGTTTGGACCATCTTAACAGACTTTTAAATTTGGCTGACC 6306
Qy 6001 TCCAAATTTATTTCCAAAGATGCGCGGCTCCCTTGTCTCTGCTCAAAAGGGGTCAAG 6060
Db 6307 TCCAAATTTATTTCCAAAGATGCGCGGCTCCCTTGTCTCTGCTCAAAAGGGGTCAAG 6366
Qy 6061 GCGGTGTGGGCGGCACTGGCATCATGACCAACAGTGTCTTTCGCGCGCAATATCTCT 6120
Db 6367 GCGGTGTGGGCGGCACTGGATCATGACCAACAGTGTCTTTCGCGCGCAATATCTCT 6426
Qy 6121 GGCATATGTCGCTTGGGCTTCATGAGAAATCAAGGGCTTAAAGCTTGCATGAATATCTG 6180
Db 6427 GGTAAATGTCGCTTGGGCTTCATGAGAAATCAAGGGCTTAAAGCTTGCATGAATATCTG 6486
Qy 6181 CAGGGGACCTTTCCTATCAATTTGTTACAGAGAGGCGCAGTGGTCCGAAACCCGCGCA 6240
Db 6487 CAGGGGACCTTTCCTATCAATTTGTTACAGAGAGGCGCAGTGGTCCGAAACCCGCGCA 6546
Qy 6241 AACTTTAAGGTCCGCATCTGAGAGTGGCGGCTCAGAGTACGCGAGGTGACGACAGAC 6300
Db 6547 AACTTAAGGTCCGCATCTGAGAGTGGCGGCTCAGAGTACGCGAGGTGACGACAGAC 6606
Qy 6301 GGGTCATACCATCACTAATACAGAGTCAACCATCTGATTAATCTTGAATATCCCTGCCAAT 6360
Db 6607 GGGTCATACCATCACTAATACAGAGTCAACCATCTGATTAATCTTGAATATCCCTGCCAAT 6666
Qy 6361 CCGTCTCCGAGTCTTTTCTGAGTGAAGTCTGCTGCTGAGTGGGCTTAAATCATTTGTCTG 6420
Db 6667 CCGTCTCCGAGTCTTTTCTGAGTGAAGTCTGCTGCTGAGTGGGCTTAAATCATTTGTCTG 6726
Qy 6421 CCGAAGCGTTTTTCCGGAGTGAAGTCTGCTGCTGAGTGGGCTTAAATCATTTGTCTG 6480
Db 6727 CCGAAGCGTTTTTCCGGAGTGAAGTCTGCTGCTGAGTGGGCTTAAATCATTTGTCTG 6786
Qy 6481 GGGTCCAGTCTCTTGGCAGCTTGAACCCGCAACAGAGATTAATGTCCATGCTTAACA 6540
Db 6787 GGGTCCAGTCTCTTGGCAGCTTGAACCCGCAACAGAGATTAATGTCCATGCTTAACA 6846
Qy 6541 GATCATCTCATATCAACGAGAGATGAGCTGAGCGGCGGCTTGAAGCGGGGATCAACCCCA 6600
Db 6847 GATCATCTCATATCAACGAGAGATGAGCTGAGCGGCGGCGGCTTGAAGCGGGGATCAACCCCA 6906
Qy 6601 TCCGAGCGAAGCTCTTCCGAGCAGCTATCCGCAACATCCGCTGCGAGCCACTGACCC 6660
Db 6907 TCCGAGCGAAGCTCTTCCGAGCAGCTATCCGCAACATCCGCTGCGAGCCACTGACCC 6966
Qy 6661 ACCCAAGCGCAAGCTTATGATGTGAGCATGTGTGGAATGCTTAACCTGTTCAATGGGGGCGAT 6720
Db 6967 ACCCAAGCGCAAGCTTATGATGTGAGCATGTGTGGAATGCTTAACCTGTTCAATGGGGGCGAT 7026
Qy 6721 GTGACTCGATATAGTCTGGGTCCAAAGTGTGCTTGTGGAATCTCTGCAACCAATGTGTC 6780
Db 7027 GTGACTCGATATAGTCTGGGTCCAAAGTGTGCTTGTGGAATCTCTGCAACCAATGTGTC 7086
Qy 6781 GAAGAAAGAGCAGCTTGAAGCTTGCATACATCAGAAATACATGTCTCCCAAGAAAGAG 6840
Db 7087 GAAGAAAGAGCAGCTTGAAGCTTGCATACATCAGAAATACATGTCTCCCAAGAAAGAG 7146
Qy 6841 TTCCCAAGCTTTACCGGCTTGGGCAAGGCTTGAATTAACCAACCAAGCTTGTGGAATG 6900
Db 7147 TTCCCAAGCTTTACCGGCTTGGGCAAGGCTTGAATTAACCAACCAAGCTTGTGGAATG 7206
Qy 6901 TGGAAAGAGCAGATTAACCAACGCGCACTGTTGCGGGCTGTGCTCTCTCTCTAG 6960
Db 7207 TGGAAAGAGCAGATTAACCAACGCGCACTGTTGCGGGCTGTGCTCTCTCTCTCTAG 7266

OY	6961	AAAAACCCCGAGCGCTCTCCCCCAAGAGAGCGCCCGGACAGTGGGCGCTTAAGTGAAGACTCTCATA	7020
Db	7267	AAGGCCCCCGAGCGCCCCCCCCCAAGAGAGAGCGCGGACAGTGGGCTTAAGGAGAGACCATTA	7326
OY	7021	GGAGATGCGCTTCAACAGCTGGCCATTAAAGTCTTTGGCCAGGCCCCCCCAAGCGGCGAT	7080
Db	7327	GGAGAGCGCTTCCAAAGCTGGCCATTAAAGTCTTTGGCCAGGCCCCCCCAAGCGGCGAT	7386
OY	7081	TCAGGCCCTTTCACCGGGGCGGGCGCTGCGGATTCGCGCACTCAGACGCCCTCTGATGAG	7140
Db	7387	TCAGGCCCTTTCACCGGGGCGGGAGCGCGCGACTTCGCGCACTCGGCTGCGGCCCTGATGAG	7446
OY	7141	TTGGCCCTTTGGAGACAGGTTCCATCTCTTTCATGCCCCCTTCAGAGGGGAGACTTGA	7200
Db	7447	TTGGCTCTTTGGAGAACAGGTTCTACCTCTCTCAGTGGCCCCCTTCAGAGGGGAGACTGGG	7506
OY	7201	GATCCAGACTGGAGCTGAGCAGGTAGAGCCCAACCCCCCCCCAGGGGGGGGTGGCA	7260
Db	7507	GATCCAGACTGGAGCTGAGCAGGTAGAGCCCAACCCCCCCCCAGGGGGGGGAGGTG	7566
OY	7261	GCTCCGCGCTCGGACTCGGGGTCTGGTCTACTTGTCTCGAGAGAGACGACTCGTGGT	7320
Db	7567	GCTCCGCGCTCGGACTCGGGGTCTGGTCTACTTGTCTCGAGAGAGAAATGTCTCGTGGT	7626
OY	7321	TGCTGCTCATGTCACTACTCTCGAGCGGGGCTCTAATACTCCTTGTAGTCCGAAGAG	7380
Db	7627	TTCTGCTCATATTCATCTCTCGAGCGGGGGGTCTAATACTCCTTGTATCTCCGAAAAA	7686
OY	7381	GAGAAGTACCGGATTAACCCCTGTGAGCAACTCCGTTGGGATGATATCAACAAGGTGAC	7440
Db	7687	GAAAGTTTCCATTTTAACCCCTTAAGCAACTCCGTGGCATACGACAAACAAGGTGAC	7746
OY	7441	TGTACCAACAAGAGCGGCTCACTAAGGCTAAAAAGTAACTTTGTATGATGATGCA	7500
Db	7747	TGTACTACATCAAGAGCGGCTCACTAGGGCTAAAAAGTAACTTTGTATGATGATGCA	7806
OY	7501	GTGCTCGACTCTACTACGACTCAGTCTTAAGACATTAAAGTACGAGGCTCCAGAGTC	7560
Db	7807	GTGCTCGAGGCTTATTATGACTCAGTCTTAAGACATCAAGCTAGCGGCTCCAAAGTTC	7866
OY	7561	AACCGAAGGCTCTCAACATGAGAGAGGCTTGCCAGTTAAACCCACCCCATTTCTGCAGA	7620
Db	7867	AGCGCAAGGCTCTCAACCTTGAAGAGGCGTGCATTTGACTCCACCCCATTTCTGCAGA	7926
OY	7621	TCTAAATATGAGGTTTGGGCTTAAGAGGTCGCGAGCTGTCCGGGAGGGCGGTTAACAC	7680
Db	7927	TCCAGATATGAGGTTTGGGCTTAAGAGGTTGCGAGCTGTTCGGGAGGGCGGTTAACAC	7986
OY	7681	ATCAAGTCCGTGTGAGAGACCTCTCTGAGACTCAGAAACACCAATTTCCCAACCAT	7740
Db	7987	ATCAAGTCCGTGTGAGAGACCTCTCTGAGACTCAGAAACACCAATTTCCCAACCATC	8046
OY	7741	ATGCGCAAAATAGAGTGTCTTGCGTGGACCCCAACAAAGGGGGCGAAGAAAGCGTTCG	7800
Db	8047	ATGCGCAAAATAGAGTGTCTTGCGTGGACCCCAACAAAGGGGGGTAAAGAAAGCGTTCG	8106
OY	7801	CTTATATGTTTACCCCTGACCTCGGCGGTCAAGGTCGCGAAGATGAGCCCTTATGACATT	7860
Db	8107	CTTATATGTTTACCCCTGACCTCGGCGGTCAAGATTTGCGAAGATGAGCCCTTATGATGTC	8166
OY	7861	ACACAAAAACTTCTCAGGCGGAGTAGGGGGCTTCTATGGAATTCAGATATTCGCCGCT	7920
Db	8167	ACACAAAAAGCTTCTCAGGCGGAGTAGGGGGCTTCTATGGAATTCAGATATTCGCCGCT	8226
OY	7921	CAGCGGGTATAGTTTCTCTTGAAGACATGGGCGGAAAAAGACCCCTATGAGGTTTTCG	7980
Db	8227	CAGCGGGTATAGTTTCTCTTGAAGCAATGGGCGGAAAAAAGACCCCTATGAGGTTTTCG	8286
OY	7981	TATGATACCCAGTGTGTTGACTCAACCGTCACTGAGAGAGACATCAAGACTGAGAGATCC	8040
Db	8287	TATGATACCCAGTGTGTTGACTCAACCGTCACTGAGAGAGACATCAAGACTGAGAGATCC	8346
OY	8041	ATATATCGGGCGTGTCTTGGCCCGAGAGGCCACACTGCACTGCTAATCTGAG	8100

[illegible]

LOCUS	AF169002	9661 bp	RNA	linear	VRL 06-SEP-2001
DEFINITION	Hepatitis C virus isolate NDM228	polyprotein gene, complete cds.			
ACCESSION	AF169002.1	GI:6707279			
VERSION	AF169002.1	GI:6707279			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
CDS					

Query Match	Best Local Similarity	Score	DB	Length
8293; Conservative	91.1%	7858.6	14	9661
Matches		Pred. No. 0	Mismatches 765	Indels 0
		Gaps 0		
ORIGIN				
ATGAGCACAATCTCTTAACCTCAAAAGAAAAACCAAAAGAAACACCAACCGTCCGCCACAA	60			
ATGAGCACAANTCTTAATCTCAAGAAAAACCAAAAGAAACCACTAACCGTCCGCCACAA	400			
GACGTTAAGTTTCCGGGCGGGCGGCGCAGATCGTTGGCGAGATATCTTTGTTCCGCGGAGG	120			
GACGTTAAGTTTCCGGGCGGGCGGCGCAGATCGTTGGCGAGATATCTTTGTTCCGCGGAGG	460			
GGCCCCAAGTTGGGTGTGGCGCGCAAGAAAGAACTTCGAGCGGTCCAGCCACGTAGG	520			
AGGCGCAAGCCATCCCTTAAAGATCGGCGCTCTCACTGGCAATCTTGAGGAAAAACAGAA	240			
AGGCGCAAGCCATCCCTTAAAGATCGGCGCTCTCACTGGCAATCTTGAGGAAAAACAGAA	580			
TACCCCTGGCCCCCTATACGGGAATGAGGAATCGGCTGGGCGAGATGGCTTCGTCCTCCC	300			
TACCCCTGGCCCCCTATACGGGAATGAGGAATCGGCTGGGCGAGATGGCTTCGTCCTCCC	640			
CGAGGTTCCCGCTCCTCTTGGGGGCCCCCAATGACCCCCCGGCAATAGTGGCCCAACGTAGG	360			
CGAGGTTCCCGCTCCTCTTGGGGGCCCCCAATGACCCCCCGGCAATAGTGGCCCAACGTAGG	700			
AAGGTCATCGATACCTTAACGTCGCGCTTTGCCGACTCATGGGGTACATCCCTGTCGTG	420			
AAAGTCATCGACACCTTAACGTCGCGCTTTGCCGACTCATGGGGTACATCCCTGTCGTG	760			
GGGCCCCCGCTGGGCGGCGGTGCGCAGAGCTTCGGGATGGCGTGAAGTCTTGGAGGAC	480			
GGGCCCCCGCTGGGCGGCGGTGCGCAGAGCTTCGGGATGGCGTGAAGTCTTGGAGGAC	820			
GGGTTTATTTTGGCAACAGGGAATCTACCTGTTGCTCTTTTCTATCTTCTTCTGCGCC	540			
GGGTTTATTTTGGCAACAGGGAATCTACCTGTTGCTCTTTTCTATCTTCTTCTGCGCC	880			
CTGCTGTCCTGCAATCCACCCCGGCTCTCGCTGCGCAAGTGAAGAACATCAAGTACCGGC	600			
CTGCTGTCCTGCAATCCACCCCGGCTCTCGCTGCGCAAGTGAAGAACATCAAGTACCGGC	940			
TACATGGATCTAATGCACTGACCTGCTCCGCTCCGCTCCCAAGTGAAGAACATCAAGTACCGGC	660			
TACATGGATCTAATGCACTGACCTGCTCCGCTCCGCTCCCAAGTGAAGAACATCAAGTACCGGC	1000			
CTCAGCGTCCCGGGGTCGTCCTGTCGGAAGATGGGGAATGATCATTCAGTCTGATAT	720			
CTCAGCGTCCCGGGGTCGTCCTGTCGGAAGATGGGGAATGATCATTCAGTCTGATAT	1060			
CCGCTTCAACCGAATGTCGCTGTCACAGCGCGCGCGCTCAAGCAAGGCTTGGAGG	780			
CCGCTTCAACCGAATGTCGCTGTCACAGCGCGCGCGCTCAAGCAAGGCTTGGAGG	1120			

QY	781	CACATCGACATGGTTGTGTAATGTCCGCCACGCTGTCTGTGCTTACGTGGGGAGACTTC	840
Db	1121	CACCTMGACATGGTTGTGATGTGCGGCACGCTCTGTTCTGCTTATAGTGGAGACCTTC	1180
QY	841	TGCGGTGGGGGTGATGTCTGCAGAGCCCAATGTATTCTTGTCTGCGCGCAGACCACTAGTTTT	900
Db	1181	TGCGCGGGGGGTGATGTCTGCAGAGCCCAATGTATTCTTGTCTGCGCGGAAACATCACTGGTTTT	1240
QY	901	GTCCAAGACTGCAATTGTCTCCATCTTACCCTGGTACCATCACTGACACACCGACATGGCATGG	960
Db	1241	GTGCAAGATGCAATTGTCTCCATCTTACCCTGGTGCATCACTGGGCAAGATATGGCATGG	1300
QY	961	GACATGATGATGAATCTGGTTCGCCGCCACGGCTACCATGATCTTGGGCTACGCCATGTGCTTC	1020
Db	1301	GACATGATGATGAATCTGGTTCGCCGCCACGGCTACCATGATCTTGGGCTATGGCATGTGCGCTTT	1360
QY	1021	CCCGAGGTCATATTATGATCATCTTATGGCGGGGCTCATTTGGGGCGTCAATTTGGGCTTTGGCC	1080
Db	1361	CCCGAGGTCATATGATCATCTTATGACATCATGACGGGGCTCATCTGGGGCGTCAATTTGGGCTTTAGCC	1420
QY	1081	TACTTCTCTATGCAAGGAGCGTGGCGAAAGTCGTTGTATCTCTTCTGTTGGCCGCCGGG	1140
Db	1421	TATTTCTTCATGACAGGAGCGGTGGCGAAAGTCGTTGTATCTCTTCTATCTGGCGGCTGGG	1480
QY	1141	GTGACGCGCGCACTTCATCTGTTGGGGGTTCTGCGCGCAGACCACTGGCGGCTCAAC	1200
Db	1481	GTGACGCGGAAATCCCTCMCTRTTYYGGGGGTCTGCTCGCGCACCAACGCTTACGAGTGTCAAC	1540
QY	1201	AGCTTATTTGACATGGGGCCCGAGAGAAATCCAGCTTCGTTAACACCAATGGCAGCTGG	1260
Db	1541	AGCTCTTCAGTATYYGGGCGCTCAGAGAAACATCCAGCTTATTAAACCAATGGCAGTTGG	1600
QY	1261	CACATCAACCGCACCGCCCTGTGAATCTGCAATGACTCTTGGACACACCGGCTTTATCGGCTCT	1320
Db	1601	CACATCAACCGCACCTGCGCTGTGAATTTGCAATGACTCTTGGAACTGTGCTTCTGCGGACC	1660
QY	1321	CTGTTCTTACACCCACAGCTTCAACTGTCAGAGATGTCCGAACGATGTCCGCTGCGCG	1380
Db	1661	CTRTTCTTACGKAAACCGCTTCAACTCGTCAAGGATGTCTTGAAACGCTGTCCGTTTGGCGCG	1720
QY	1381	AGTATCGAGGCTTTCGGGTGGGATGGGGCGCTTGGCAATATGAGGATATATGTCAACAT	1440
Db	1721	AATATCGAGGCTTTCGGGATAGGAATGGGGCACCTTGGCAATATCGAGGATATATGTCAACAT	1780
QY	1441	CGAAGATATGAGACCCCTATTTGTGTGGGACTAACCCACCAAGGAGGTGGGTGTCTCC	1500
Db	1781	CGAAGGATATGAGACCAATTTGCTGGGACTAACCCACCAAGGATATGATGTATGTCTCC	1840
QY	1501	GCGAAGACTGTGTGTGGCCAGGTGTACTGTTTCAACCCCAAGCCAGTGTGTGTGGGACG	1560
Db	1841	GCGAGGTCTGTGTGTGGCCCGGTGTGTACTGTTTCAACCCCAAGTGTGTGTGTGGGACG	1900
QY	1561	ACCGACAGGCTTTGAGGCGCCCACTTACACGTGGGGGGAGAAATGAGACAGATGTCTTCTTA	1620
Db	1901	ACCGATACACGTGAGGTCTTACTTACACATGGGGAGGAATGAGACAGATGTCTTCTTA	1960
QY	1621	TTGAACAGACACTCGACACACGCTGGGGGTCATGGTTCCGCTGACAGTGAATGAATCTTTCT	1680
Db	1961	CTGAACAGACACCCGACACACCGAGGGGTCATGGTTCCGCTGACAGTGAATGAATCTTCACT	2020
QY	1681	GAGTACACCAAGACTTGTGGGGGCAACACCTGTGCTTACAGCTGACTTTCAGCGCCAGC	1740
Db	2021	GAGTATACCAAGACTTGTGGGGGCAACACCTGTGTGTACCAAGCTGACTTTCAGTCCAGC	2080
QY	1741	ACGGAAGCTGTTGTGCCCCACGGAAGCTGTTTATGGAAGACATCTGATATCACTTAACTCTAA	1800
Db	2081	ACTGACCTGTTGTGCCCCACGGAAGCTGTTTATGGAAGACATCTGGAAGCACTTAACTCTAA	2140
QY	1801	TGCGGCTCTGGGGCCCTGGCTACACGCGCAAGGTGCGTGAATGCACTTACCCCTACAGGCTCGG	1860
Db	2141	TGTGCTTCTGGGCTTGTGGCTGACGCGCAAGGTGCGTGTGAATTACCCCTTACAGGCTCTCGG	2200
QY	1861	CATTACCCCTGCACAGTTAACTATTAACATCTTCAAAATAGAGATGATGTGGAGGGGTTT	1920

[illegible]

3281 AGTCGATGAGAGAGAAAGTTATCGTCTGGGGAGCGAGACGCGCTGACATGCGGGGACATC 3340
3001 TTACACGAGACTCCCGGTGCGCGCCGACCTGGTGGAGAGTCTGCTTGGCCAGCTGAT 3060
3341 TTACACGAGACTTCTGTGTCCGCCGCTCGGTGGAGAGATCTCTTGGCCAGCTGAT 3400
3061 GGCTATACCTCCAGAGGGGTGAGTCTTCTCGCCCCATCACTGCTTACGCCACAGACA 3120
3401 GGCTACACCTCCAGAGGGGTGAGGCTTCTTGGCCCCATCACGCTTACGCCACAGACA 3460
3121 CGTGACCTTTGGGACCAATAGTGTGAGCAATGAGGGGGCGGACAAACAGAAACAGCT 3180
3461 CAGGATCTTGGGGTGTATAGTGTGAGTATGACGGGGCGGACAAACAGAAACAGGCC 3520
3181 GGGGAAATTCAGAGTCTGTCCACAGTCACTAGTCTTCTCGAAATCCATCTCGGGG 3240
3521 GGGGAGATCCAGTCTGTCCACGCTCACTAGTCTTCTTGGAAATCCATCTCGGGG 3580
3241 GTTTGTGAGCTGTCTACCATGAGAGCTGGCAACAGACTCTGGCCGCTCACGGGGTCCG 3300
3581 GTCTATGAGCTGTCTACCAACGAGAGCTGGCAATAGACTTACGCGGCTCACGGGGCCA 3640
3301 GTACGAGAGATGACTCCAGTGTGAGGGGGGACTTAAAGGTGGCCAGGCCCTCCGGG 3360
3641 GTACCCAGATGTACTCCAGTGTGAGGGGGGAGCTTGTAGGGTGGCCAGGCCCTCTGGA 3700
3361 ACTAATCTTTGAGACCGGTGACAGTGTGAGAGCGGTGACCTGTACCTGGTACGCGGAC 3420
3701 ACTAATCATTTGAGACCGGTGACAGTGTGAGAGCGGTGACCTGTACCTGGTACGCGGAC 3760
3421 GCTGATGTCATCCCGCTCGAAGACGCGGGGCAAAACGGGGAGCGCTACTCTCCCGAGA 3480
3761 GCTGACGTATCCCGGCTCGAAGACGCGGGGCAAAACGGGGAGCGCTTCTCTCCCGAGA 3820
3481 CCTCTTCCACCTTGAAGGGGGTCTCAAGGGCCGGGTGCTATGCCCCAGGGGGCAACGCT 3540
3821 CCTCTTCCACCTTGAAGGGGGTCTCAAGGGGGAGCGGTGCTTGTGCCAGGGGGCAACGCT 3880
3541 GTCGGAGTCTTCGCGGACAGCTGTGTCTCTCGGGGCGTGGCTAAGTCCATAGATTTTC 3600
3881 GTCGGGATCTTCGCGGACAGCTGTGTCTCTCGGGGCGTGGCTAAGTCCATAGATTTTC 3940
3601 CCGGTGAGACACTGACATCGTCAAGCGGTCCCCACCTTTAGTGAACAAGACACACA 3660
3941 CCGTGTGAGACACTGACATCGTCAAGCGGTCTCCACCTTTAGTGAACAAGACACACA 4000
3661 CCGTGTGAGCGCCAGACCTATCAGGTGGGTACTTGCATGCCCGGACCTGGGAGTGAAG 3720
4001 CCGCTGTGCGCCAGACCTATCAGGTGGGTATTTGCATGCCCGGACCTGGGAGTGAAG 4060
3721 AGCACCAAGTTCCTGTGCATATGCTGTCAAGGGGTATTAAGTGTAGTGTCTTATCC 3780
4061 AGCACCAAGTTCCTGTGCATATGCTGTCAAGGGGTATTAAGTGTAGTGTCTTATCC 4120
3781 TTAGTGGCTGCGACCTGTGGGTTTGGGGGTACTTGTCTAAGGCAATGGCATCAATCC 3840
4121 TCGGTGGCTGCGACCTGTGGGTTTGGGGGTACTTGTCTAAGGCAATGGCATCAATCC 4180
3841 AACATTAGGACTGAGTCAAGACTGTGACGACCGGGGGGCCATACGTAATCCCATAT 3900
4181 AACATTAGGACTGAGTCAAGACTGTGACGACCGGGGGGCCATACGTAATCCCATAT 4240
3901 GGCAAATTCCTGCGCGATGGGGGCTGTGCGGGCGGCTAGACATCATATATGAT 3960
4241 GGCAAATTCCTGCGCGATGGGGGCTGTGCGGGCGGCTAGACATCATATATGAT 4300
3961 GAATGCGATGCGGTGAGCTTACACACATCTTGGCATGGAAACAGTCTTGATCAAGA 4020
4301 GAATGCGATGCGGTGAGCTTACTACTATCTCGGCATCGAAACAGTCTTGATCAAGA 4360
4021 GAGACAGCTGGGGTCAAGACTATCTGTGCTGCTAGACGTAAGCGCCCTGGGGTCAAGCA 4080
4361 GAGACAGCTGGGGTCAAGACTATCTGTGCTGCTAGACGTAAGCGCCCTGGGGTCAAGCA 4420

4081 ACCCCCAACCCCAATTAGAGAGAGTGCCCTTGGCAGAGAGCGAGATCCCTTCTAT 4140
4421 ACCCCCAATCCCAATTAGAGAGAGTAGCCCTTAGGCGACGAGGGTGAATCCCTTCTAT 4480
4141 GGGAGGGCGATTCCTGTCTTACATCAAGAGAGAAACATGTATCTTCTGCATTTCA 4200
4481 GGGAGGGCGATTCCTGTCTTACATCAAGAGAGAAAGCACATGTATCTTCTGTCACTCA 4540
4201 AAGAAAAGTGAAGAGCTCGGGGGGCGCTTGGGGGTATGGGCTTGAATCAGTGGCA 4260
4541 AAGAAAAGTGAAGAGCTCTGGGGCGCTTAGGGGCAATGGGCTTGAACGCTGGGCA 4600
4261 TACTAGAGAGGGTGGACGCTTCCGTAAATCAACACTCAGGAGACGTAAGTGTCTGCGC 4320
4601 TATTACAGAGGGTGGACGCTTCCGCAATACAGCTCAGGAGAGATGATGAGTGTGCTGCGC 4660
4321 ACCGAGCGCTCATGACAGGGTATCTGGGAGACTTGGATCTCGGTATGCACTGCAAGTGA 4380
4661 ACCGAGCGCTCATGACAGGGGTATCTGGAGACTTGGATCTCGGTATGCACTGCAAGTGA 4720
4381 GCGGTCACTCAAGTTGAGACTTCAGTTTGAACCCCAATTCACATTAACACAGATT 4440
4721 GCGGTCACTCAAGTTGAGACTTCAGCTTAGACTTAGACCCCACTTACCATTAACACAGACC 4780
4441 GTCCTCAAGACGCTGTCTCAAGTACAGCGCCCGGGGTGCAACGGGTAGGGAGAGACTG 4500
4781 GTCCTCAAGACGCTGTCTCAAGTACAGCGCCCGGGGTGCAAGGAGAGAGAGACTG 4840
4501 GGCATTTATAGTATGTTTTCATCTGAGACAGGCTCAGAAATGTTTGAACAGTGAAGT 4560
4841 GGCATTTATAGTATGTTTTCATCTGAGACAGGCTCAGAAATGTTTGAACAGGTAAGT 4900
4561 CTCGTGAGTGTACAGACGAGGGGCGCATGGTATGAGTCAACCATCGAGACACC 4620
4901 CTCGTGAGTGTACAGACGAGGGGCGCATGGTATGAGTCAACCATCGAGACACC 4960
4621 GTCAGGCTCAGGGCGTATTTCAACAGCGCCGGTTTGCCTGTGTGCCAAGACATCTTGA 4680
4961 GTCAGGCTCAGGCGATATTTCAACAGCGCGCTGCGCTGCGGTATGCCAAGTCACTTGA 5020
4681 TTTTGGAGGAGTTCACCGGCTTCACACATAGATGCCACTCTTCTTCCCAACA 4740
5021 TTTTGGAGGAGTTCACCGGCTTCACACATAGATGCCACTCTTCTTCCCAACA 5080
4741 AAGCAATCGGGGAAATTTGCACTTAAGCCTTACAGGCTACAGGTGCGCTAGG 4800
5081 AAGCAATCGGGGAAATTTGCACTTAAGCCTTACAGGCTACAGGTGCGCTAGG 5140
4801 GCCAAGCCCCCCCCGCTTGGGACGTATGTGAAGTGTGACTGCACTCAAGCCC 4860
5141 GCCAAGCTCTCTCGCTGCTGGAGGTGATGGAAGGTGACCCGACTCAAGCCC 5200
4861 ACACTGTGGGCCCCACACTCTCTCTGTACCGCTTGGGCTCTGTTACCAACGAGGTAC 4920
5201 ACCGTTGTGGGCCCCACACTCTCTGTGTACCGTTGGGCTCTGTTACCAACGAAAGTACC 5260
4921 CTCACACATCCGTTGAGAAATACATGCGCACTGTGATGCAAGCGACCTTGAAGTCA 4980
5261 CTCACACATCTTGAAGAAATACATGCGCACTGTGATGCAAGCTATCTTGAAGTCA 5320
4981 ACCAGCATGGGTCTTGGCAGGGGAGTCTTGGCGGCGGTGCGCGATTTGCTGGCG 5040
5321 ACCAGCATGGGTCTTGGCAGGGGAGTCTTGGCGGCGGTGCGCGATTTGCTGGCG 5380
5041 ACCGGGTGTGTTTGCATCATCGGCGCTTGCATATTAACAGAGAACCGGTCTTGGCGCG 5100
5381 ACCGGGTGTGTTTGCATCATCGGCGCTTGCATATTAACAGAGAACCGGTCTTGGCGCG 5440
5101 GACAAAGAGTCCCTATATGAGGCTTTTGAAGATGAGGAATGTGCTCTTGAAGCGGCT 5160
5441 GACAAAGAGTCCCTATATGAGGCTTTTGAAGATGAGGAATGTGCTCTTGAAGCGGCT 5500

OY	5161	CTCATTTGAAGAGGGGCGAGCGGATATAGCCGAATGCTGAAAGTCCAGATTCAGAGGCTTATTTG	5220
Db	5501	CTCATTTGAAGAGGGGCGAGCGGATATAGCTGAATGCTGAAGTCCAAATTTCAAGGCTTATTTG	5560
OY	5221	CAGCAAGGCTTCCAAACAGACTCAAGACATTAACAACCACTGTGCAGGGCTTACTTGGCCCAAG	5280
Db	5561	CAGCAAGGCTTCCAAACAGAGCCGAGACATTAACCCCGTGTTCAGAGCTTCAATGGCCCAAG	5620
OY	5281	GTAGAAACAATTCTGGGCCCAACACATGTGAACTTCAATTAGCGGATCCAAATACCTGCGA	5340
Db	5621	GTGGAGCAATTTCTGGGCGCAAGCACATGTGAACTTCAATTAGCGGATCCAAATACCTGCGG	5680
OY	5341	GGACTATTAACAACATCCGCGAGGAAACCCCTGCACTAGTCTTCAGATGTGGCTTCAATGTCCGC	5400
Db	5681	GGACTGTAAACGTCCGCGAGAAACCCCGCTGTGCTTCCATGAATGTGCATTTCAATGTCCGC	5740
OY	5401	CTCACACAGTCCGCTGTCAACAAGCACACACTATCTCTTCAACATTTTGGGGGGCTGGGCTA	5460
Db	5741	CTCACACAGCCCGTGTCAACTAGCACCACTATCTCTTAAACATCTGTGGAGGCTGGGCTG	5800
OY	5461	GCATTCGCAAAATTTGACACACCCGCGGGGCGCACTGGCTTCTGTTGTCAATGTGCGCTAGTGGGA	5520
Db	5801	GGGTCCCAATATTTGGCGCACCCGCGGGGCGCACCGGCTTCTGTCTGATGTGCTGTGGGG	5860
OY	5521	GCTGCGGTAGGCACTATATGCGCTTAGAGTAAAGTGTCTAGTGACATCTGCGAGGGTATGCT	5580
Db	5861	GCTGCGGTAGGCGAGGTATGAGGCTTGGGGCAAGGTGTGCTGATGTGACATCTGCGAGGGTATGCG	5920
OY	5581	GGGGGCAATTTGGGGGGGCTCTGTGCGCAATTCAGAAATCATGTCTGGCGAGAAACCCCTCCATG	5640
Db	5921	GGGGGCAATTTGGGGGGGCGCTGTGTGCAATTCAGAGTCAATGTCTGGCGAGAAACCCCTCCATG	5980
OY	5641	GAGAGTGTCTGCAAACTTGTCTGCTGCTGAGAAATTTCTGTCTCCGGGTGCTTGGTATGGAGTCT	5700
Db	5981	GAAATGTCTGCAAACTGTCTGTGCTGCTGAGAAATTTCTGTCTCCGGGTGCTTGGTATGGAGTCT	6040
OY	5701	ATCTGCGCGGCGCAATTTGTGCGCGCGACACGTGGGACCGGGGGAAGGCGCGCTTCCAAATGATG	5760
Db	6041	ATCTGCGCGGCGCAATCTGTGCGCGCGCACGTGGGGCCGGGGGAAGGCGCGCTTCCAAATGATG	6100
OY	5761	AATAGACTCATTTGGCTTTGGCTTCCAGAGAAATACGTCGCGCCCCACCCACTTAAGTGACG	5820
Db	6101	AACAGGCTCATTTGGCTTTCGCTTCCAGAGAAACACAGTTGGCCCCACCCACTTAAGTGACG	6160
OY	5821	GAGTCGAGTGTGCTGCGACGCTGTGACCCCAACTACTTTGGCTTCCACTTAACAGACCTG	5880
Db	6161	GAGTCGAGTGTGCTGCGACGCTGTGACCCAGCTGCTTGGCTTCCACTTAACAGACCTG	6220
OY	5881	CTCAGAAAGACTCCAACACTGTGATTACTGAGGACTGCCCCATCCCATGTGGCGGCGCTGTGG	5940
Db	6221	CTCAGAAAGACTCCAATTAATTTGATTACTGAGGACTGTGCCATCCCATGTGGCGGCGCTGTGG	6280
OY	5941	CTTCGCGCATGTGTGGGACTGGGTTTGCACCATCTCTTAACAGCTTTAAAAATTGGCTGAC	6000
Db	6281	CTTCGCGCATGTGTGGGACTGTGGGTTTGCACCATCTCTTAACAGCTTTAAAAATTGGCTGAC	6340
OY	6001	TCCAATTAATTTCCAAAGATGCGCGGCGCTCCCTTTGTCTCTGTGCAAAAAGGGGTAACAG	6060
Db	6341	TCCAATTTGTTTCCAAAGATGCGCGGCGCTCCCTTTATCTTTGTGCAAAAAGGGGTAATAGA	6400
OY	6061	GGCGTGTGGCGCGGACTGTGGCATATGACACACGATGTCTTTGCGGCGCGCAATATCTCT	6120
Db	6401	GGCGTATGTGGCGCGGACTGTGGCATATGACACACGATGTCTTTGCGGCGCGCAATATCTCT	6460
OY	6121	GGCAATGTCTGGCTTGGGCTCCATGAGAAATCACGGGGCTTAAAGCTGTGATGAATATCTGG	6180
Db	6461	GGCAATGTCTGGGCTTGGGCTCTTATGAGATCAACGGGCTTAAAGCTGTGATGAATATCTGG	6520
OY	6181	CAGGGGACCTTTCCATCAATTTGTTTACACGAGGGGCGAGTCTGTGCGAAACCCGCGCGCA	6240
Db	6521	CAGGGGACCTTTCCATCAATTTGTTTACACGAGGGGCGAGTCTGTGCGAAACCCGCGCGCA	6580
OY	6241	AACTTTAAGTCTGCCATCTGAGAGGTGCGGCTTCAGATTCGCGGAGTGAACGACGAC	6300

Db	6581	AACTTCAAACCCGCACTCTGGAGAGTGGCGGCTTCGGAGTACGGGAGGTGACACAGAC	6640
Oy	6301	GGGTCATACCACTACATACAGAGACTCCACACTGATTAACCTTGAAGAGTCCCTGCGCAACTA	6360
Db	6641	GGGCGCTACTCTATATTAACAGAGACTGACACTGACAAACCTGAAGAGTTCCCTGTCAATA	6700
Oy	6361	CCCTCTCCGAGTTCTTTTCTGGGTGAGCGAAGTGCAGATCCATAGGTTTTCGCCACA	6420
Db	6701	CCCTCTCCAGAGTTCTTTTCCGTGGGTGAGCGAAGTGCMAATCCATAGGTTTCTCCCAAG	6760
Oy	6421	CCGAAGCCGTTTTTCCGGAGTAGAGGTCTCGTTCGCGTGGGCTTAATCACTTGTGCTC	6480
Db	6761	CCAAAGCCGTTTTTCCGGAGTAGAGGTCTCGTTCGCGTGGGCTTAATCACTTGTGCTC	6820
Oy	6481	GGGTCCCACTCTCTTGGGAGCCCTGAACCCGACACAGACGTATTGATGTCCATGCTMA	6540
Db	6821	GGGTCTCAGCTTCTCTTGTGAGCCCTGAACCCGACACAGACGTATTGATGTCCATGCTMA	6880
Oy	6541	GATCCATCTATATCAACGCGGAGACTGACGCGGCGGCTTTAGCGCGGGGCTCACCCCCA	6600
Db	6881	GATCCATCTCCATATTAACGCGGAGGCTGACACACGCGCTTAGCGCGGGGCTCACCCCCA	6940
Oy	6601	TCCGAGGCAACTCTCTCGGCGAGCCAGTATCGGACCATGCTGCGGACCACTTGACCC	6660
Db	6941	TCTGAGGCAACTCTCTCAGCAAGCAGCTGTGCGACACATGCTGCGGCGCACTGTGCC	7000
Oy	6661	ACCCAGGCAAGCCATATGATGTGACATGTGTGATGCTMACTGTTCATGGGGGCGAT	6720
Db	7001	ACCCAGGCAAGCCATATGATGTGACATGTGTGATGCTMACTGTTCATGGGGGCGAT	7060
Oy	6721	GTGACTCGGATAGAGTCTGGGTCCAAAGTGTGCTTGTGACTCTCTGACCCATGATC	6780
Db	7061	GTGACTCGGATAGAGTCTGAGTCCAAAGTGTGCTTGTGACTCTCTGACCCATGATC	7120
Oy	6781	GAAAGAAAGAGCGACTTGAAGCTTCGATACCATCAGAAATACATGCTTCCCAAGAAAGG	6840
Db	7121	GAAAGAAAGAGCGACTCTGAGACCTTCGATACCATCAGAAATGATGCTTCCCAAGAAAGG	7180
Oy	6841	TTTCCACAGCTTAAACGGGCTGGGCAACGGCTGATTAACAACCCACGCTTGTGAAATCG	6900
Db	7181	TTTCCACAGCTTAAACGGGCTGGGCAACGGCTGATTAACAACCCCTTGTGAAATCG	7240
Oy	6901	TGGAAGAGGCGAGATTACCAACCGGCGCACTGTTGCGGCTGTGCTTCCCTCTTAGG	6960
Db	7241	TGGAAGAGGCGCGATTACCAACCGGCGCACTGTTGCGGCTGTGCTTCCCTCTTAGG	7300
Oy	6961	AAAAACCCGAGCGCTCCCCCAAGAGGCGCGGACAGTGGCTTAAGTAGAATCCATA	7020
Db	7301	AAGGCTTCGAGCGCCCCCCCAAGAGAGCTGTGGACGGGTGTTGAGTAGAGACCATTA	7360
Oy	7021	GGAGATGCGCTTCAAACAGCTGGCCATTAAGTCTTTGGCCAGCCCCCCCCCAAGGGGCGAT	7080
Db	7361	GGAGATGCTCTTCAAACAGCTGGCCATTAAGTCTTTGGCCAGCCCCCCCCCAAGGGGCGAC	7420
Oy	7081	TCAAGGCTTTTACAGGGGGGCGGCGCTGCGCATTTCCGCAAGTCAAGAGCTCTCTGATAG	7140
Db	7421	TCAAGGCTTTTACAGGGGGGCGGAGCTGTGCGAATCTCGGCGGCGAGGCGTCCCGGATAG	7480
Oy	7141	TTGGCCCTTTTGGAGACAGTTTCATCTTTTCATGCCCCCTTCGAGGGGAGACTTGA	7200
Db	7481	TTGGCTTTTGGAGAGACAGTTTCTAACCTCTTCATGCCCCCTTCGAGGGGAGACTTGGG	7540
Oy	7201	GATCCAGACCTGAGAGCTGAGACAGTAGAGGCCCAACCCCCCGGAGGGGGGCTGCA	7260
Db	7541	GATCCAGACCTGAGAGCTGAGACAGTAGAGTTCAGAGGTCTTCTCCACAGGAGGAGGAGTGC	7600
Oy	7261	GCTCCCGGCTCGGAGCTCGGGGTCTCGGTCTACTTGTCTCCGAGAGAGAGCACTCGTGTG	7320
Db	7601	GCTCCCGGCTCGGAGCTCGGGGTCTCGGTCTACTTGTCTCCGAGAGAGAGCACTCTGTGTG	7660
Oy	7321	TGCTGCTCATGTCAATCTCTTGAACGGGGCTTAATTAATCTCTTGTAGTCCGAAAG	7380

Db 7661 TGCTGCTCATGTCTACTCTCTGAGACCGGGGCTCTAATACTCTTCCGGGCCGGAAG 7720
 Qy 7281 GAGAGATTACCGATTAAACCCCTTGAGCAACTCCCTGTGGCATATCACAAGAAGTGTAC 7440
 Db 7721 GAGAGATTGCGCAATCAACCCCTTGAGCAACTCTGCTGTGGCATATCACAAGAAGTGTAC 7780
 Qy 7441 TGTACCAACAAGAGAGCGCTCACTAAGAGGCTAAAAAGGTAACTTTTGTAGTAGTACAA 7500
 Db 7781 TGTACTACATCAAGAGACCGCTCACTAGAGGCTAAAAAGGTAACTTTTGTAGTAGTACAA 7840
 Qy 7501 GTGCTGCACTCTCTACTAGACTAGTCTTAAAGAGCATTAAGCTAAGGCGCTCCAAAGTC 7560
 Db 7841 GTGCTGCAAGCGCTAATTAAGACTCAACTTAAAGAGCATTAAGCTAAGGCGCTCCAAAGTC 7900
 Qy 7561 ACCGCAAGGCTCTCACCATAGAGAGGCTTGCCAGTTAACTCCAGCCCAATTCGTGCAAGA 7620
 Db 7901 AGCGCAAGGCTCTCTTACTTGAAGAGAGCGTGTAGTAACTCCAGCCCAATTCGTGCAAGA 7960
 Qy 7621 TCTAATAATAGGCTTTGGGGCTTAAGAGAGTCCGAGCTTTGTCGGGAGGCGCTTAACAC 7680
 Db 7961 TCCAGATACGGGTTTGGGGCTTAAGAGAGTCCGAGCTTTGTCGGGAGGCGCTTAACAC 8020
 Qy 7681 ATCAAGTCCGCTGAGAGAGACTCTGAGAGAGCTCAAGAAACCAATTTCCCAACACAT 7740
 Db 8021 ATCAAGTCCGCTGAGAGAGACTCTGAGAGAGCTCAAGAAACCAATTTCCCAACACAT 8080
 Qy 7741 ATGACCAAAAAATGAGTGTCTGCTGAGAGACCCCAACAAGGGGGCAAGAAAGCAGTCGC 7800
 Db 8081 ATGACCAAAAAATGAGTGTCTGCTGAGAGACCCCAACAAGGGGGTAAAGAAAGCAGTCGC 8140
 Qy 7801 CTTATCGTTTAACTTACCTGACCTGCGCTCAAGGCTTGCGAAGAAAGATGCGCTTTATGACAT 7860
 Db 8141 CTTATCGTTTAACTTACCTGACCTGCGCTCAAGGCTTGCGAAGAAAGATGCGCTTTATGACAT 8200
 Qy 7861 ACAAAAAAATCTTCTCAGCGGCTGATGCGGGCTCTTATGAGATTCACAGTATCCCGCGCT 7920
 Db 8201 ACAAAAAAATCTTCTCAGCGGCTGATGCGGGCTCTTATGAGATTCACAGTATCCCGCGCT 8260
 Qy 7921 CAGCGGGTAGAGTTCTCTGAAAGCATGAGCGGAGAAAGAGAGACCTTATGAGTATTCG 7980
 Db 8261 CAGCGGGTAGAGTTCTCTGAAAGCATGAGCGGAGAAAGAGAGACCTTATGAGTATTCG 8320
 Qy 7981 TATGATACCCGATGCTTTGACTCAACCGTCACTGAGAGAGACATGAGAGCTGAGAGATCC 8040
 Db 8321 TATGATACCCGATGCTTTGACTCAACCGTCACTGAGAGAGACATGAGAGCTGAGAGATCC 8380
 Qy 8041 ATATATGCGGCTGCTCTCTGCGAGAGAGGCGCACTGCGATACATCACTGCTTATGAG 8100
 Db 8381 ATATATGCGGCTGCTCTCTGCGAGAGAGGCGCGCACTGCGATACATCACTGCTTATGAG 8440
 Qy 8101 AGACTTTACGTGAGAGGCGCTTATTTCAACAGAAAGGCGCAAACTTGGGGTACAGGCGT 8160
 Db 8441 AGACTTTACGTAGGTGGGCGCATGCTGATATGCAAGGGCGCAAGGCTCGGGTACAGGCGT 8500
 Qy 8161 TGCGCGCGCAGCGGGGTGCTCAACATGAGATGAGGAAACACCATCACTGCTTATGAG 8220
 Db 8501 TGCGCGCGCAGCGGGGTGCTCAACATGAGATGAGGAAACACCATCACTGCTTATGAG 8280
 Qy 8221 GCCTTAGCGCTTGTAAAGCTGAGAGGATATGCGGCGCAAGTGTGATATGCGGAGT 8280
 Db 8561 GCCTTAGCGCTTGTAAAGCTGAGAGGATATGCGGCGCAAGTGTGATATGCGGAGT 8620
 Qy 8281 GACTTGTGTGCTCTAGAAAGCCAGAGGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAG 8340
 Db 8621 GACTTGTGTGCTCTAGAAAGCCAGAGGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAG 8680
 Qy 8341 TTCAAGAGGCTATGAGAGAGATTTCTGCGCTCTGCTGAGAGAGAGAGAGAGAGAGAT 8400
 Db 8681 TTCAAGAGGCTATGAGAGAGATTTCTGCGCTCTGCTGAGAGAGAGAGAGAGAGAT 8740
 Qy 8401 GATCTGAGAGCTATGAGAGATTTCTGCGCTCTGCTGAGAGAGAGAGAGAGAGAGAT 8460
 Db 8741 GATCTGAGAGCTATGAGAGATTTCTGCGCTCTGCTGAGAGAGAGAGAGAGAGAT 8800

Qy 8461 CGCGGAGATTAATCTACCTGACAGAGAGACCTTACCACTCAATGCGCGGGCTGCTGGAA 8520
 Db 8801 CGCGGAGATTAATCTACCTGACAGAGAGACCTTACCACTCAATGCGCGGGCTGCTGGAA 8860
 Qy 8521 ACAGTTAGACACTCTCCCTGTCAATTGATGAGGAGGAGAAATCATCAAGTACGCGCGGAC 8580
 Db 8861 ACAGTTAGACACTCTCCCTGTCAATTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 8920
 Qy 8581 ATATGAGGCTGCGATGAGTCTGATGAGACACTTCTTCAATCTCATGAGTCAAGAGACAG 8640
 Db 8921 ATATGAGGCTGCGATGAGTCTGATGAGACACTTCTTCAATCTCATGAGTCAAGAGACAG 8980
 Qy 8641 CTGAGACAGAACTCAACTTTGAGATGATGAGAGCGGCTGACTCCGTGAGTCCCTTGAC 8700
 Db 8981 CTGAGACAGAACTCAACTTTGAGATGATGAGAGCGGCTGACTCCGTGAGTCCCTTGAC 9040
 Qy 8701 CTCCAGCTAATAATTGAAGGTTACATGAGGCTTGAAGCTTTTCTCTGACACATACACT 8760
 Db 9041 CTCCAGCTAATAATTGAAGGTTACATGAGGCTTGAAGCTTTTCTCTGACACATACACT 9100
 Qy 8761 CCCAGAGAACTGACAGAGGAGGCTTCAAGCCTTCAAGAACTTGGGGGCGCACCCCTCAG 8820
 Db 9101 CCCAGAGAACTGACAGAGGAGGCTTCAAGCCTTCAAGAACTTGGGGGCGCACCCCTCAG 9160
 Qy 8821 GCGTGAAGAGCGGAGCAGTGCAGTCAAGGCGTCCCTCATCTCCGTGAGGAGAGAGCG 8880
 Db 9161 GCGTGAAGAGCGGAGCAGTGCAGTCAAGGCGTCCCTCATCTCCGTGAGGAGAGAGCG 9220
 Qy 8881 GCGGTTTGGGCTGATATCTCTTCAATTGGGCGGTGAAGACCAAGTCAAACTCACTCA 8940
 Db 9221 GCGGCTGCGGCTGATATCTCTTCAACTGGGCGGTGAAGACCAAGTCAAACTCACTCA 9280
 Qy 8941 TTGCGGAAAGCGGCGCTCCGATTAATCAAGTGTATCAAGTCCGCGCGCGCGGCGGCG 9000
 Db 9281 TTGCGGAAAGCGGCGCTCCGATTAATCAAGTGTATCAAGTCCGCGCGCGCGGCGGCG 9340
 Qy 9001 GACATTTATCACAGCGTGTGCGGTCGCCGACCCCGCTTATTTGCTTTGGCTCACTCTCA 9060
 Db 9341 GACATTTATCACAGCGTGTGCGGTCGCCGACCCCGCTTATTTGCTTTGGCTCACTCTCA 9400
 Qy 9061 CTTTTTGTAGGGGTAGGCTTTTCTTCACTCCCGGCTCGG 9099
 Db 9401 CTTTTTGTAGGGGTAGGCTTTTCTTCACTCCCGGCTCGG 9439

RESULT 15

AF238483 LOCUS 9416 bp RNA linear VRL 26-MAR-2000
 Hepatitis C virus 2a polypeptide gene, complete cds.

AF238483
 AF238483
 AF238483.1 GI:7329204

SOURCE

Hepatitis C virus
 Hepatitis C virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE

1 (bases 1 to 9416)
 Itakura,J., Nagayama,K., Enomoto,N., Kurosaki,M., Watanabe,H. and Sato,C.

TITLE

Full length cDNA sequence of HCV genotype 2a, strain MD2a-4
 Unpublished

JOURNAL

Submitted (24-FEB-2000) Second Department of Internal Medicine,
 Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-Ku,
 Tokyo 113-8519, Japan

FEATURES
 source 1..9416
 /organism="Hepatitis C virus"
 /mol_type="genomic RNA"

CDS

```
/strain="MD2a-4"
/db xref="taxon:11103"
/notice="genotype: 2a"
307..9408
/codon_start=1
/product="polyprotein"
/protein_id="AF59942.1"
/db xref="GI:7329205"
/translation="MSTNPKPQKRTKNTNRPRQDYKPGGQIVGVVILPRRPL
GVARTKTSERQPRGRQPIPKRSTGSKGPGYPMFLVGLGMLGLSLPRG
SRPSWPTDPRHRSRVNGKVIDPLTGCFLDMGYIIPVAGPLGAVRTAAHGVLE
GVNATGNTLPGCSFSLFLLALSCITIPVAIEVNTSDYNTNCSNISTMOJLA
AVLHVGCPVCEKQNTSRMIPSPVAVQARGATLQGRITIDIMVNSATCSALY
VGLDGGVMLAAMPIYTPQHMVFQDNCISYIGTITGRMADMMNMSPATML
AVMARPEVILIDITGAMVMPGLAFVNSQGMARVYVILLTAGVDANTYVSGST
AHLQTPAGLFTLSQONIQILNTSGMHNRTALNCESLNTFLASLTHTAFNS
GCEPRLSACNIEAFRIGMGLQYEEVNTNEDMRPCHMYPKQCGIIPARVSCGV
YCEPSPVVGTIDKLGVPYTWGENEDVFLNSTPRPGAMFTGTFKTC
GAPCRIRADFNASTDLCTPCFRKPEATYKCGSPMLTFRGMVDPYRGIHPC
TNNYTI EKIRMYGVGBHLTAQNFTRGRCNLEDRSGLSLHSTTEMILLPCT
FSLDPLASTGLHLHONIVDYQYMGSLPATIRIYRMENVVILLILADARCACTM
MLTLLEQAEALEKLVLAHASASCSGFLFYVFAVAYIIGRAAPLALYLTGLM
SFLCLLALPEQAVAYADASHQIGAVLLVTLFTLPQKILSLRCLMWLYLTLL
GEMIDBWPYPMQVGRDGI IWAATIFCGVVFDTKMLALRGLYLRGALTREP
YFVARAHALLMCTMARHLAGKXYOMALLAGWGTGYIDHLTPMSDMAAGLRDLA
VAVEPIIESPEKKVIVMGAEFAACGDIHLGLPVSARLGEVILGSPADVTSRGMKL
APITAAQOTRGLGAI VSMTRGDKTEQAGEIQULSTYQSLGTSISVLTMYHG
AGKTLTAGSRPVTQMSASGDLVGPSPGRTKSLBCTCGAVDLVLTAKADVIPA
RRGDRGRGALLSPPLSTLKSSGGPVLCPRHAGVIFRAVCSRGAKSIDIPET
LDIVTRSPTFSDNSTPAVPQIYOVGLAHPTGSKSTKPVAAVAAQGYVLLINB
GATLFGQAYLSKAYGNPNIRTVRTYTGDSITYSGFLADGGCAGADYI IICD
ECHAAGSTNMLIGTVLDOAETAGVRLTVATATLPGSVETPHPNIEVALGQGEIP
FYGRAPLSTIKGRHILFCHSKKCDLAAALRGMANAVAYRGLDVSIVITQGV
VVVAIDLMATGTFGDEPTVIDCNVAVTQVDFSLDPFTITITVTPQDAVRSQRK
TGRGLGTVYVSTGERASGDFSVLCECYDAGAAVETLPEATTYRLAYNTPEL
PVCODHLEFWEAVFTGLIHDAHFLSQTQAGSPFLVAYQATVCARAPPSMDI
MKCITRELKPLVGPFLYRLGTSVTEITLTPVTYKILTCQADLEVTSTVMI
GIAVAAYACLATGVCSIIIGRLINORAIYAPDKVLYEAFDEMEBASQIILBEO
RIAEMLKSKIQGLLOQASKOADIOPAVQASPRKVPFAKHMNPTSGIYALRST
LPGNPVASMMAFSAALTSPLSTSTILLNLGRMWLSQIAPAGAPFVAGSLVGA
VGNIGLQKLYVDLIAGAGISGALVAFKMSGKPSMEVILMLPLBPGALVGV
ICAILLRHVGPAGAVQMMNRLIAFASRKNHVAPTHYVESASQVDTLGLSLIT
SLRLHNMITTEDCPIPCAGSLMDWDWCTIITDPMKSLPKPMKMGIFISQ
KGYKWAGCIGMITTRPCOGANSVNRVLSMRITGKTCOMIWOCPPTNCTGEOC
VPRPAPFKTAIRVASEYAEATYHQQYSYVTGLTDDNLKPCQLPSEPFSSWDY
QIHRFAPTPPFRFDEVSFYVGLNSFVVSQLPCEPDPDVLMSMLTESHITAEAA
ARLIARGSPSEBASSSQSLASPLRATCTHGAVDVMDVNDLFWGDDVTRIEES
KVVLVSLDPMAEKSDLEPSISEVMLPNRPALPAPARDYNPPLVESMKRPDY
QPTTVGALPPKPTPTPPRRRTYVGLDESTIGALQOLAIKTGQIIPASDGLS
TGAADSGRTSPDELALSENSISGMPRLBGBRPDLPEPQVQLQPPQGEIAP
GSDSGASTCESBDDVCCSMYSWTGAILTPCOSPEBEKLPINLSNLSLRHKNKY
CTTSKASLSLAKRVTFDRQVLDAAHYDSVKDIKLASKAPSLNLTLERACQLPHS
ARSKYFGAEKERSLSGRANVHITSWKDLLEDSQPIPTIIMAKNEVFCVDPTEKX
KARLLIVPLDGRVCEKALYDITOKLPQAVGASGFIYSAORFPLMAAEK
DPGFSYDTRCFDSVTERDIRTEBSIYRASKPEBHATHSLTELYLVGMINSK
GSGCYRCAAGSGLVLTSMGNTITCYKALACKAGVIA PMLVAGSDI VLVSEQ
TEEDERNLRAFTEAMTRYSAFPEDPREPEDELITSCSNVVALNPLRERYILTR
DPTTPIARAAMETVRHSVNSWLGNI IQVAPTIWARVNLWTHFSLIMADDTQINL
FENYGVSVSPLDPAIERHGLDAFSLHYSPHLTVAVTLGAPLPRAMKS
RAVAVRASLISRGRAVCGRYLFNMAVVKLKLTPLEARLILDSMFWVAGGGDI
YHSVSRARPLRLGLLLSVGVFLPLPAR"
```

ORIGIN

Query Match 86.3%; Score 7852.2; DB 14; Length 9416;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 8319; Conservative 0; Mismatches 780; Indels 0; Gaps 0;

```
QY 1 ATGAGCACAATCTTAACCTCAAGAAAAACCAAAAGAAACACCAACCGTGGCCGCA 60
D 307 ATGAGCACAATCTTAACCTCAAGAAAAACCAAAAGAAACACCAACCGTGGCCGCA 366
QY 61 GACGTTAAGTTTCCGGCGCGCGCAGATCGTTGGCGAGTATACTTGTGGCCGCGAG 120
```

```
D 367 GACGTTAAGTTTCCGGCGCGCGCAGATCGTTGGCGAGTATACTTGTGGCCGCGAG 426
QY 121 GGGCCGAGGTTGGGTGTGCGCGGACAAAGAAACATTCCGAGGGTCCGACCGTGA 180
D 427 GGGCCGAGGTTGGGTGTGCGCGGACAAAGAAACATTCCGAGGGTCCGACCGTGA 486
QY 181 AGGCGCAGGCCATCTCCCTAAAGATCGGCGCTCCATCGGCAATCTCGGGAAAAACAGA 240
D 487 AGGCGCAGGCCATCTCCCAAGATCGGCGCTCCATCGGCAAGTCTCGGGAAAAACAGA 546
QY 241 TACCCCTGGCCCTATACGGGAATGAGGACTCGGCTGGCAGAGATGCTCTGCC 300
D 547 TACCCCTGGCCCTATATAGGAATGAGGAGCTCGGCTGGCAGAGATGCTCTGCC 606
QY 301 CGAGTTCCTGCTCCCTCTTGGGGGCCCCAATGACCCCGCATAGTTCGCGCAACGTGG 360
D 607 CGAGTTCCTGCTCCCTCTTGGGGGCCCCAATGACCCCGCATAGTTCGCGCAACGTGG 666
QY 361 AAGGTATCGATACCTTAACGTCGCGCTTTCGCACTCATYGGGGTACATCCCTGCTGTG 420
D 667 AAGGTATCGATACCTTAACGTCGCGCTTTCGCACTCATYGGGGTACATCCCTGCTGTG 726
QY 421 GGGCGCCGCTCGGCGGCGCTCGCAGAGCTTCGCGCATGCGGTGAGAGTCTGAGAGAC 480
D 727 GGGCGCCGCTCGGCGGCGCTCGCAGAGCTTCGCGCATGCGGTGAGAGTCTGAGAGAC 786
QY 481 GGGGTATATTTTGCACAGGGAACCTTACCCTGTTGCTCTTTCTATCTTCTGCTGGCC 540
D 787 GGGGTATATTTTGCACAGGGAACCTTACCCTGTTGCTCTTTCTATCTTCTGCTGGCC 846
QY 541 CTGCTGTCTGCATCAACACCCCGCTCCGCTGCGGAGTGAAGTGAAGAACATCAATCCGG 600
D 847 CTACTGTCTGCATCAACATCTCCGCTCTCTGCTATGGAATGAAGAACATCAATCCGG 906
QY 601 TACATGATGATCAAGCATGCGACCAATGACAGATTAAGTTCGAGCTTCAGGCTGCTG 660
D 907 TATATGATGATCAAGCATGCTTCAATGAGCATCACTCGGAACTCCAGGCGCGCTG 966
QY 661 CTGCAAGTCCCGGGGTGCGTCCGCTCGAGAAAGTGGGGAATGCATCTCATGCTGGATA 720
D 967 CTGCAAGTCCCGGGGTGCGTCCGCTCGAGAAAGTGGGGAATGCATCTCATGCTGGATA 1026
QY 721 CCGGTCTCAACGAAATGCGCGTGCAGCGGCGCGGCGCTCAGCAGGCGTTGCGGAGC 780
D 1027 CCGGTCTCAACGAAATGCGCGTGCAGCGGCGCGGCGCTCAGCAGGCGTTGCGGAGC 1086
QY 781 CACATGACATGATGATGATGTCGCGCAGCTGCTGCTGCTTACGTTGCGGAGACTC 840
D 1087 CACATGACATGATGATGATGTCGCGCAGCTGCTGCTGCTTACGTTGCGGAGACTC 1146
QY 841 TCGGATGGGGTATGCTCGCAGCGCAAAATGTTATTTCTCGCGCGAGACACTGTTT 900
D 1147 TCGGATGGGGTATGCTCGCAGCGCAAAATGTTATTTCTCGCGCGAGACACTGTTT 1206
QY 901 GTCGAAGATGCAATGCTCATCTACCTGATACCATCACTGGAACACCGCATGCGATG 960
D 1207 GTCGAAGATGCAATGCTCATCTACCTGATACCATCACTGGAACACCGCATGCGATG 1266
QY 961 GACATGATGATGATGATGTCGCGCAGCGCTACCATGATCTTGGCGTACGCGATGCTG 1020
D 1267 GACATGATGATGATGATGTCGCGCAGCGCTACCATGATCTTGGCGTACGCGATGCTG 1326
QY 1021 CCGGAGGTCTTATACATCATTAACGCGGCTCATTTGGGGCGTCAATGTTCCGCTGGCC 1080
D 1327 CCGGAGGTCTTATACATCATTAACGCGGCTCATTTGGGGCGTCAATGTTCCGCTGGCC 1386
QY 1081 TACTTCTATGACAGGAGCGTGGGCGAAAGTGGTGTCACTCTTGTGTCGCGCGG 1140
D 1387 TACTTCTATGACAGGAGCGTGGGCGAAAGTGGTGTCACTCTTGTGTCGCGCGG 1446
QY 1141 GTGAGCGCGGCAACCATACTGTTGGGGGTTCTGCGCGGAGACACCGGCGCTCAC 1200
D 1447 GTGAGCGCGAATACCTACTCGTGGGGGTTCCATGCGCATCTCACCGACCTTGTGCC 1506
```

OY 1201 AGCTTATTGACATGAGGCCGAGAGAAATCCAGCTCGTTTACACCAATGGCAGCTGG 1260
DB 1507 GGCCTCTTACCCCTGGGCTCTCAGCAGAAATTCAGCTCATACCAATGGCAGTTGG 1566
OY 1261 CACATCAACGCGACCGCCCTGAACTGCATGACTCTCTTGCACACCGGCTTTATCGCTCT 1320
DB 1567 CACATCAACCGCTACTGCGCTGAATTTGCATGAATCCCTGAAACACCGGCTTTCTCGCTCC 1526
OY 1321 CTGTTTACACCGACAGCTTCACTCGTCAGAGATGTCGCCGAAAGCATGTCGCGCTGCGC 1380
DB 1627 CTGTTTACACCGACAGCTTCACTCGTCAGAGATGTCGCCGAAAGCATGTCGCGCTGCGC 1686
OY 1381 AGTATCGAGGCTTCCGGGTGGAGTGGGCGCCTTGGCAATATGAGATTAATGTCACCAAT 1440
DB 1687 AACATCGAGGCTTCCGGATAGAGATGGGCGCTTTTGGCAATACGAGAAATGTCACCAAT 1746
OY 1441 CCAGAGATATGAGACCTTATTTGCTGGCACTACCCACCAAGGCAATGTCGCTGCTCC 1500
DB 1747 CCAGAGATATGAGACCTTATTTGCTGGCACTACCCACCAAGGCAATGTCGCTGCTCC 1806
OY 1501 GCGAAGACTGTGTGGGCGGAGTAACTGTTTCAACCCCGACGAGTGAATGGGAGCG 1560
DB 1807 GCGAGTCTGTATGTGGCGCGGTACTGTTTCACTCCAGCCGAGTGTGGGAGCG 1866
OY 1561 ACCGACAGGCTTGGAGCGCCCACTTACAGTGGGGGAGAAATGAGACAGATGTCCTCTA 1620
DB 1867 ACCGATAGCTTGGAGTGGCTTATATACATGGGGAAGAAATGAGACAGATGTCCTCTA 1926
OY 1621 TTGAACAGACTCGACACCGCTGGGTGATGTGCTGGCTGCACGTGATGAATCTTCT 1680
DB 1927 TTGAATAGCACCCGGCCACCGCGGGGGCATGTTCGGCTGCAAGTGAATGAATCTCCACT 1986
OY 1681 GGCATACACAGACTTGGGGGGGACACACCTGCGCTACTAGAGTGAATCTTCAAGCGAGC 1740
DB 1987 GGTTCACCAAGACTTGGGGGACACACCTTGGCGCATAGAGTGAATCTTCAAGCGAGC 2046
OY 1741 ACCGACCTGTGTGGCGCCGACGAGCTGTTTATAGAAAGCATCTGATATACACTTACCTGAAA 1800
DB 2047 ACCGACCTGTGTATGCGCCACCGACTGTTTATAGAAAGCATCTGATATACACTTACCTGAAA 2106
OY 1801 TGGCGCTTGGGCGCTTGGCTGACACGCAAGTGGCTGATGACTACCTTACAGGCTTGG 1860
DB 2107 TGTGGCTTGGGCGCTTGGCTGACACGCAAGTGGCTGATGACTACCTTACAGGCTTGG 2166
OY 1861 CATTAACCCGTCACAGTTAATCATTAACATCTTCAAAATAGAGATATGTTGGAGGGTT 1920
DB 2167 ATTCATCTCTGACAGTCACTAATCAATCTTCAAGATTAAGATGTTGGGGGGGTT 2226
OY 1921 GAGCAGAGGCTCAGCGCTGACATGCAATTTCACTCGTGGGAGTCTTGGCACTTGGAGGAC 1980
DB 2227 GAGCAGAGGCTTACGGCGCGCGGTGTAATCTTCACTCGTGGGAGTCTTGGCACTTGGAGGAC 2286
OY 1981 AGAGACAGAGTCAAATGTCTCTCTTGTGTCATCTCCACCGAGAAATGGGCAATTTTAACT 2040
DB 2287 AGAGACAGAGTCAAATGTCTCTCTTGTGTCATCTCCACCGAGAAATGGGCAATTTTAACT 2346
OY 2041 TGTCTTACTCGAGACCTGCGCGCTGTGTCAGTGTCTTCCACTCCCTCCACCAAAACATC 2100
DB 2347 TGTACTTCTCAGACTTGCCTGCTTGTGACTGTCTTCTCCACTCCCTCCACCAAAACATC 2406
OY 2101 GTGACGTACAATTCATGATATGCTATACCTGCGCTCAGCAAAATACATGTCGAGTGG 2160
DB 2407 GTGACGTGCAATATATGATATGCTATACCTGCGCTCAGCAAAATACATGTCGAGTGG 2466
OY 2161 GAGTGGTAAATATCTTATATCTCTGCTTTAGCGAGCGCCAGGGTTTGGCTGCTTATGG 2220
DB 2467 GAGTGGTAAATATCTTATATCTCTGCTTTAGCGAGCGCCAGGGTTTGGCTGCTTATGG 2526
OY 2221 ATGTCATCTTGTGGGCGAGCGGAGCAGACACTAAGAGAGTGTGCTATCTTGCAGGCT 2280
DB 2527 ATGTCATCTTGTGGGCGAGCGGAGCAGACACTAAGAGAGTGTGCTATCTTGCAGGCT 2586

OY 2281 GCGAGCGAGCTAGCTGCAATGAGCTTCTATATTTTGTATCTTTTCTGCTGCTGTTGG 2340
DB 2587 GCGAGCGAGCTAGCTGCAATGAGCTTCTATATTTTGTATCTTTCTGCTGCTGCTG 2646
OY 2341 TACATCAAGGCTGGGTAGTCCCTTAACTATTTCCCTCACTGAGGCTGTGTCTTT 2400
DB 2647 TACATCAAGGCGGCGAGTCCCTTGTGCTATTTCCCTTAAACGCGCTATGTCTTTC 2706
OY 2401 AGCTACTGCTCTATAGACTTGGCCCAACAGGCTTATGCTTATAGACGATCTGTGATGAC 2460
DB 2707 TGTCTACTGCTCTATAGACTTGGCCCAACAGGCTTATGCTTATAGACGATCTGTGATGAC 2766
OY 2461 CAGATAGAGCGGCTGTGCTGTATGATGATCACTCTTTTACTCTACCCCGAGTATAAG 2520
DB 2767 CAGATAGAGCGGCTTGTGCTGTATGATGATGATGATGATGATGATGATGATGATGATG 2826
OY 2521 ACCCTTCTCAGCGGTTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
DB 2827 ATCTTCTTAAAGCGGCTTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2886
OY 2581 GTCACAGAGTGGGACACCACTATGACAGTGGCGGCGGTGTGTGTGTGTGTGTGTGTGTGTGT 2640
DB 2887 ATTCAGAGATGGGTACCACTCATGACAGTGGCGGCGGTGTGTGTGTGTGTGTGTGTGTGTGT 2946
OY 2641 GTGCGCATATTTCAACCAAGT 2700
DB 2947 GCCACCATATTTCTGCGCGGT 3006
OY 2701 GGGCTGTCTTACTCTTAAAGGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2760
DB 3007 GGGCTGTGTATCTCTTAAAGGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3066
OY 2761 GCTCTACTGAGATGTGCAACATGCAAGGCAATCTGCGGGGGGCGAGTACGTCAGATG 2820
DB 3067 GCTCTCTGAGATGTGCAACATGCAAGGCAATCTGCGGGGGGCGAGTACGTCAGATG 3126
OY 2821 GCGCTACTAGCTTGGCAGGTGAGTGGCACTTATCATATGATCACTACCTCATG 2880
DB 3127 GCGTATATAGCTTGGT 3186
OY 2881 TCGGATTTGGCTGTAGTGGCTTGGCGGACCTTGGCGGTGTGTGTGTGTGTGTGTGTGTGTGT 2940
DB 3187 TCGGATTTGGCTGTAGTGGCTTGGCGGACCTTGGCGGTGTGTGTGTGTGTGTGTGTGTGTGT 3246
OY 2941 AGTCGATGGAAGAAAGTCAATGTCTGAGGAGCGGAGACAGCTCTTGTGGGAGCAT 3000
DB 3247 AGTCGATGGAAGAAAGTCAATGTCTGAGGAGCGGAGACAGCTCTTGTGGGAGCAT 3306
OY 3001 TTACACGACCTTCCGCTGTCCGCGCACTTGTGTGGAGGCTCTCTTGTGGCCGAGTAT 3060
DB 3307 CTACACGACCTTCCGCTGTCCGCGCACTTGTGTGGAGGCTCTCTTGTGGCCGAGTAT 3366
OY 3061 GGCATTAATCTTCAAGGGGTGAGTCTTCTGCGCCCATCATGTGCTTAAAGCCAGAGACA 3120
DB 3367 GGCATTAATCTTCAAGGGGTGAGTCTTCTGCGCCCATCATGTGCTTAAAGCCAGAGACA 3426
OY 3121 CGTGGCTTTTGGGACCAATAGTGTGAGATGACGAGGAGCGGAGCAAGAACAGAGCT 3180
DB 3427 CGAGTCTCTTGGGTGCTATGATGTGATGATGATGATGATGATGATGATGATGATGATGATG 3486
OY 3181 GGGGAAATTCAGGCTCTGTGTCACAGTCACTGAGTCTTCTCTGGAACATTCATCTCGGG 3240
DB 3487 GGGGAAATTCAGGCTCTGTGTCACAGTCACTGAGTCTTCTCTGGAACATTCATTCAGGG 3546
OY 3241 GTTTTGTGAGCTGTCTTACATGAGGCTGTGGAACAACAATCTTGGCGGCTTACAGGGGTCCG 3300
DB 3547 GTCTTATGAGCTGTCTTACATGAGGCTGTGGAACAACAATCTTGGCGGCTTACAGGGGTCCG 3506
OY 3301 GTCACGAGATGATCTCAGTGTGCTGAGGGGAGCTTATAGAGTGGCCAGCCCTCTGGG 3360
DB 3607 GTCACGAGATGATCTCAGTGTGCTGAGGGGAGCTTATAGAGTGGCCAGCCCTCTGGG 3666
OY 3361 ACTAAATCTTTGAGCGGTGACAGTGTGAGCGGTGACCTGTATCTGTGCTACGCGGAC 3420

Db 3667 ACCAAATCTTTGGAGCGGTGACGTGTGAGCGGTCCACCTGTACCTGTGTAACGCGGAGC 3726
Qy 3421 GCTGATGTCAATCCCGGCTCCGAAGACGGGGGACAAACGGGGAGCGCTACTCTCCCGAGA 3480
Db 3727 GCTGATGTCAATCCCGGCTCCGAAGACGGGGGACAAACGGGGAGCGGTACTCTCCCGAGA 3786
Qy 3481 CCTCTTCCACCTTGAAGGGGTCTCAAGAGCGCGGTCGTAATGCCCAAGGGGACAGCT 3540
Db 3787 CCCCCTTCAACCTTGAAGGGGTCTCAAGAGCGCGGTCGTAATGCCCAAGGGGACAGCT 3846
Qy 3541 GTCGAGCTCTTCCGGGAGCGTGTGTCTCTCGGGGCGTGTGCTAATGATTTTCATC 3600
Db 3847 GTGGGAGTCTTCCGGGAGCGTGTGTCTCTCGGGGCGTGTGCTAATGATTTTCATC 3906
Qy 3601 CCCGTTAAGACACTCGACATGTGTACCGGGTCCCACTTTAGTGAACAAGACACCA 3660
Db 3907 CCCGTTAAGACACTGTGTACCGGGTCTCCACCTTTAGTGAACAAGACACCA 3966
Qy 3661 CCTGCTGTGCCCCGACCTTACAGGTGGGGTACTTGTGATGCCCGACGTGGAGTGAAG 3720
Db 3967 CCGGCTGTGCCCCGACCTTACAGGTGGGGTACTTGTGATGCCCGACGTGGAGTGAAG 4026
Qy 3721 AGCACCAAGTTCCTGTGTGATATGTCTCAAGGGGTATAAGTGTAGTCTTAATCCC 3780
Db 4027 AGCACCAAGTTCCTGTGTGATATGTCTCAAGGGGTATAAGTGTAGTCTTAATCCC 4086
Qy 3781 TCAGTGTGCTGACCTCTGGGGTTGGGGCGTACTTGTCTAAGGACATGGCATATCCC 3840
Db 4087 TCGGTGGGTGCTCACTTGTGGGGTTGGGGCGTACTTGTCTAAGGACATGGCATATCCC 4146
Qy 3841 AACATTAGGACTGGAGTCAAGGACTGTGACGACCGGGGCGCCATCAGTACTCCACATAT 3900
Db 4147 AACATTAGGACTGGAGTCAAGGACTGTGACGACCGGGGCGCCATCAGTACTCCACATAT 4206
Qy 3901 GGGAAATTCCTCGCCGATGTGGGGGCTGTGCGGGCGGCGCTACGACATCATATGTGAT 3960
Db 4207 GGTAAATTCCTCGCTGATGTGGGGGCTGTGCGGGCGGCGCTACGACATCATATGTGAT 4266
Qy 3961 GAATGCCATCCGTGTGATCTAACCAATCTGGGTATCGGAACAGTCTTGTATCAAGCA 4020
Db 4267 GAATGCCATCCGTGTGATCTAACCAATCTGGGTATCGGAACAGTCTTGTATCAAGCA 4326
Qy 4021 GAACAGCTGGGGGTCAAGCTTAATGTGTGTGCTTACACTACGCCCCCTGGGTCAATGACA 4080
Db 4327 GAACAGCTGGGGGTCAAGCTTAATGTGTGTGCTTACACTACGCCCCCTGGGTCAATGACA 4386
Qy 4081 ACCCCCCACCCCAACATAGAGAGGTGGCCCTTGGGAGAGAGGCGAGATCCCTTCTAT 4140
Db 4387 ACCCCCCATCCCAATATAGAGAGGTGGCCCTTGGGAGAGAGGAGGTGAGATCCCTTCTAT 4446
Qy 4141 GGGAGGGCGAATCCCTGTCTTATCATCAAGGGAGGAAGACATGTATCTTCTGCAATCA 4200
Db 4447 GGGAGGGCGAATCCCTGTCTTATCATCAAGGGAGGGGACCTTGTATCTTCTGCAATCA 4506
Qy 4201 AAGAAAAGTGTACGAGCTCGGGGCGGCTTCCGGGTATGGGCTTGAATCTCAATGAGA 4260
Db 4507 AAGAAAAGTGTACGAGCTCGGGGCGGCTTCCGGGTATGGGCTTGAATCTCAATGAGA 4566
Qy 4261 TACTACAGAGGGTGTGACGTCTCCGTAAATACAACTCAAGGGAGAGCTAGTGTGTGCGC 4320
Db 4567 TACTACAGAGGGTGTGACGTCTCCGTAAATACAACTCAAGGGAGAGTATAGTGTGTGCGC 4626
Qy 4321 ACCGAGCGCCTCATGACAGGGTATATCTGGGAGCTTTGATCTCGTATGACCTGACAGTA 4380
Db 4627 ACCGAGCGCCTCATGACAGGGTATATCTGGGAGCTTTGATCTCGTATGACCTGACAGTA 4686
Qy 4381 GCGGTCACTCAAGTGTGAGCTTCAAGTTTGAACCCCACTTCAACATTAACACAGATAT 4440
Db 4687 GCGGTCACTCAAGTGTGAGCTTCAAGTTTGAACCCCACTTCAACATTAACACAGATAT 4746
Qy 4441 GTCCCTCAAGACGCTGTCTCAAGTACGACGCGCGGGGTGCGACGAGGTAGGGAGAGCTG 4500

Db 4747 GTACTCAAGACGCTGTCTCAAGTACGACGAGCGCGGGGGCGACAGGGTATAGGAAGACTG 4806
Qy 4501 GGCATTTATATAGTATGTTTCCACTGTGTGAGAGAGCTTCAGGAATGTTTGAAGTGTAGTG 4560
Db 4807 GGCATTTATATAGTATGTTTCCACTGTGTGAGAGAGCTTCAGGAATGTTTGAAGTGTAGTG 4866
Qy 4561 CTCGTGATGCTACGACGAGGGGGCGCATGGTATAGGCTCAACCATCGAGACCAACC 4620
Db 4867 CTCGTGATGCTACGACGAGGGGGCGCATGGTATAGGCTCAACCATCGAGACCAACC 4926
Qy 4621 GTCAGGCTCAAGGGGTATTTCAACACGCGCGTGTGCTGTGTGTCGCAAGACATCTTGAG 4680
Db 4927 GTCAGGCTCAAGGGGTATTTCAACACGCGCGTGTGCTGTGTGTCGCAAGACATCTTGAG 4986
Qy 4681 TTTTGGAGGAGCTTTTCAACCGGCTTCAACACATATATGCTCCACTTCTTCCCAACA 4740
Db 4987 TTTTGGAGGAGCTTTTCAACCGGCTTCAACACATATATGCTCCACTTCTTCCCAACA 5046
Qy 4741 AAGCAATCGGGGGAAATTTGCACTTAAACAGCTTACAGGCTTACAGGCTTACAGGCTTACG 4800
Db 5047 AAGCAATCGGGGGAGAGAGCTTTCATACCTGTGTGCTTATCAAGCTTACAGTATGCGCTAGA 5106
Qy 4801 GCCAAAGCCCCCCCCCGTCTGTGGAGCGTATGTGAAGTGTGACTGACTCAAGCCCC 4860
Db 5107 GCCAAAGCCCCCCCCCGTCTGTGGAGCATTATGTGAAGTGTGACTTCAAGCTCAAGCCCC 5166
Qy 4861 AACTCGTGGGGCCCCACACTCTCTCTATACCGCTTGGGCTCTGTATACCAAGAGGTCAAC 4920
Db 5167 AACTCGTGGGGCCCCACACTCTCTCTATACCGCTTGGGCTCTGTATACCAAGAGGTCAAC 5226
Qy 4921 CTCACACATCCCGTGAAGGAATATCATGCGCACCTGATGGAAGGCGGACTTGTAGGTATG 4980
Db 5227 CTCACACATCCCGTGAAGGAATATCATGCGCACCTGATGGAAGGCGGACTTGTAGGTATG 5286
Qy 4981 ACCAGCAATGAGTCTTGTGGAGGGGAGTCTTGGCGGCGTGTGCGCGGTATGCTGCGG 5040
Db 5287 ACCAGCAATGAGTCTTGTGGAGGGGAGTCTTGGCGGCGTGTGCGCGGTATGCTGCGG 5346
Qy 5041 ACCGGGTGTGTGATCATGCGGCGCTTGTGACATTTAACAGAGAGCGGTCTGCGCGG 5100
Db 5347 ACCGGGTGTGTGATCATGCGGCGCTTGTGACATTTAACAGAGAGCGGTCTGCGCGG 5406
Qy 5101 GACAGGAGGTCCCTATATGAGGCTTTGATGAGTGGAGGAATGTGCTGTAGGGCGAGCT 5160
Db 5407 GACAGGAGGTCCCTATATGAGGCTTTGATGAGTGGAGGAATGTGCTGTAAAGCGAGCT 5466
Qy 5161 CTCATTGAAAGAGGGGAGCGAGATAGCCGAGATGCTGAAGTCCAAAGTCCAAAGCTTATG 5220
Db 5467 CTCATTGAAAGAGGGGAGCGAGATAGCCGAGATGCTGAAGTCCAAAGTCCAAAGCTTATG 5526
Qy 5221 CAGCAAGCTTCCAAACAGCTCAAGACATTAACCACTGTGAGGCTTCAATGCGGCAAG 5280
Db 5527 CAGCAAGCTTCCAAACAGGCGCAGGACATTAACCAAGGCTGTGAGGCTTCAATGCGGCAAG 5586
Qy 5281 GTAGAACAAATTTGGGGCAACACATGTGGAACCTTCAATGAGCGGCAATACCTGCA 5340
Db 5587 GTAGAACAAATTTGGGGCAACACATGTGGAACCTTCAATGAGCGGCAATACCTGCA 5646
Qy 5341 GAGACTATCAACTGCGAGGGAACCTGCAAGTTCATGATGCGGCTTCAAGTCCAGCGCC 5400
Db 5647 GAGACTATCAACTGCGAGGGAACCTGCGTGTGCTTCAATGATGCGGCTTCAAGTCCAGCGCC 5706
Qy 5401 CTCACAGTCCGCTGTCAACAGACATATCTTCTCAACATTTTGGGGGGCTGTGCTA 5460
Db 5707 CTCACAGTCCGCTGTCAACAGACATATCTTCTCAACATTTTGGGGAGGTGTGCTG 5766
Qy 5461 GCAATCCAAATTTGACCAACCGGGGGGCGCACTGAGCTTCTGTTCAGTGTGCTTATGAGGA 5520
Db 5767 GCAATCCAAATTTGCTCAACCGGGGGGCGCACTGAGCTTCTGTTCAGTGTGCTTATGAGGA 5826
Qy 5521 GCTGCGGTAGGAGTATAGGCTTATAGTATGCTATAGTGAACATCTGTGAGAGGTATGCT 5580
Db 5827 GCTGCGGTAGGAGTATAGGCTTATAGTATGCTATAGTGAACATCTGTGAGAGGTATGCT 5886

OY	5581	GGGGGCAATTTGGGGGGGCTCTGCTGGCAATTCAGATCAATGCTGTGGCGAAGAGCCCTTCATG	5640
Db	5887	GCGGGCAATTTGGGGGGGCCCTCTGTGCATTTCAAGATCATGCTGTGGCGAAGAGCCCTTCATG	5946
OY	5641	GAGATGTCGCAACTTGCTGCTGCGAATTTGTCTCGGGGTGGCTTGATAGTGGAGTC	5700
Db	5947	GAAATGTCATCAACCTGTGCTGCGGATCTCTGTCCCGGGTGGCCGTGTGTGGGGGTTC	6006
OY	5701	ATCTGCGGCGCAATTCGTGCGCGACAGTGGGACCGGGGGAAAGGCCGCTTCATGTGATG	5760
Db	6007	ATCTGCGGCGCAATTCGTGCGCGCACCTGTGGACCGGGGGAAAGGCCGCTTCATGTGATG	6066
OY	5761	AATAGATCATTTGGCTTTGCTTCCAGAGGAAATACAGTCGCGCCCAACCACTAGTGAAG	5820
Db	6067	AACAGGCTCATTCGCTTGTGCTTCCAGAGGAAACAGTCGCGCCCTACTCATCTAGTGAAG	6126
OY	5821	GAGTCGAGTGGGTGCGACAGCGTGTGACCCAACTAGTTGGCTCCCTTACCATTAACAGGCTG	5880
Db	6127	GAGTCAGATGGGTGCGACAGCGTGTGACCCAGCTGGCTTGCTCTTTAACAATTAAGCTTA	6186
OY	5881	CTCAGAGACTTCACAACTGGATTACTGAGGACTGCCCCCATCCCATGCGGCGCTCGTG	5940
Db	6187	CTCAGAGAACTTCACAAATTTGATTACTGAGGACTGCCCCCATCCCATGCGGCGCTCGTG	6246
OY	5941	CTCGGCGAGTGTGGGGACTGGGTTTTGACCAATCCCAACAGACTTTAAAAATTTGGCTAGCC	6000
Db	6247	CTTCGCGAGTGTGGGAGCTGGGTTTTGACCAATCTGACAGATTTTAAAACTTGCTAGCT	6306
OY	6001	TCCAAATTAATTCGCAAAAGATGCCGGGCTCCCTCTTTGTCCTCTGTCAAAAAGGGGTACAG	6060
Db	6307	TCCAAAGTTGTTCCCAAGATGCTGCTGGCTCCCTCTTTATCTTGTGCAAAAAGGGGTACAG	6366
OY	6061	GCGGTGTGGGCGGCACTGGGCAATCAACAGGTTCCTTGGCGGCGCAATATCTCT	6120
Db	6367	GGTGTGTGGGCGGCACTGGGCAATCAACAGATGCTCCCTGGGCGCAAAATTTCT	6426
OY	6121	GGCAATGCGGCTTGGGCTTCATAGAGATTCACGGGGGCTTAAGCCTGCAATGAATCTCG	6180
Db	6427	GGCAATGCTCGGCTGGGGCTCAATGAGATTCACAGGGGCTTAAGACCTGTATAAACATTTGG	6486
OY	6181	CAGGGGACCTTTCTCATATTTGTTTACACGGAGGGGCGAGTGGCCGAAACCCGCGCA	6240
Db	6487	CAGGGGACCTTTCTCATATCTATCACTGTTAACCTGAGGGGCGAGTGGCCGAAACCCGCGCG	6546
OY	6241	AACCTTAAAGGTTCGCATCTGAGAGGTGGCGGCTTCAGAGTACGGGAGGTGACCGAGCAC	6300
Db	6547	AACCTTCAAGACCGGCATCTGAGAGGTGGCGGCTTCAGAGTACGGGAGGTGACCGAGCAC	6606
OY	6301	GGGTGATACCACTATACAGGAATCAACACTGATTAACCTTGAAGAGTCCCGTCGCAACTA	6360
Db	6607	GGGCAATCTCTGATGTGACAGAGATTTGACCACTGACCACTTGAAGAGTCCCGTCGCAACTA	6666
OY	6361	CCCTCTCCCGAGTTCTTTTCTGTGGGTGACGGAGTGCAGATTCATAGGTTTTGCCCCACA	6420
Db	6667	CCCTTCTCAAGAGTTCTTTTCTGTGGGTGACGGAGTGCAGAAATCATAGGTTTTGCCCCACA	6726
OY	6421	CCGAAGCCGTTTTTCCGGGATGAGGTCTTCGCTTTCGCGTTAGTTCAATTTGTGCTC	6480
Db	6727	CCGAAGCCGTTTTTCCGGGATGAGGTCTTCGCTTTATTTGTGGGCTCAATTCGTTGTGCTTA	6786
OY	6481	GGGTCCCAAGCTTCCTTGGGACCCTGAAACCCGACAGAGGATTTGATGTCATGCTGTAACA	6540
Db	6787	GGGTCTCAAGCTTCCTCTGTGATCTTGAGCCCGATTAACAGATGTGTGATTTCCATGTTTACA	6846
OY	6541	GATTCATCTCATATACAGGCGGAGACTGACAGCGCGGCTTTTAGCGCGGGGTCAACCCCA	6600
Db	6847	GAAACATCCCAATATTACGGCGGAGGCTGAGCGCGGCTTTAGCAACGGGGGTCAACCCCG	6906
OY	6601	TCGAGGCGAAGTCTCTCGGCGAGCGAGCTATGGGCACTATGCTGTGCGAGCCACTGCAACC	6660
Db	6907	TCTGAGGCGAGCTCTTTCAGCGAGCGAGCTATCAGACCATGCTGCGAGCCAACCTGCAACC	6966

QY	6661	ACCACGGCAAAAGCTATGATGTGGACAATGGTGAAGTCAAACTGTTCAATGGGGGGGGAT	6722
Db	6367	ACCACGGCAAAAGCTTATGATGTGGACAATGGTGAAGTCAAACTGTTCAATGGGGGGGGAT	7026
QY	6721	GTGATCGGATAGAGTCTGGAGTCCAAAGTGTCGATTCGTGGAATCTCTGACCCCAATGTC	6780
Db	7027	GTGATCGGATAGAAATCTGAGTCAAAAGTGTCGATTCGTGGAATCTCTGACCCCAATGTC	7086
QY	6781	GAAATAAGAGCGGACTTGGAGCCTTGATACATACGATATACATAGCTCCCAAGAAAGG	6840
Db	7087	GAAATAAGAGCGGACTTCGAGCCTCAATACATCGAGTATATGCTCCCAAGAAAGG	7146
QY	6841	TTCCACGAGCTTACCGGCTGGGCACGGCTGATTACAAACCAACCGCTTGTGAATCG	6900
Db	7147	TTCCACGAGCTTACCGGCTGGGCACGGCTGATTACAAACCAACCGCTTGTGAATCG	7206
QY	6901	TGGAATAAGCGCAGATTACAAACCGGCACTGTTGCGGCTGTGTCTCTCCCTCCTAAG	6960
Db	7207	TGGAATAAGCGCAGATTACAAACCGCCTGTTGCGGCTGTGTCTCTCCCTCCTAAG	7266
QY	6961	AAACCCCGACGCTCCCTCCCAAGAGGCGCGGACATGAGGCTTAAGTGAAGATCCATA	7020
Db	7267	AAACCCCGACGCTCCCTCCCAAGAGGCGCGGACATGAGGCTTAAGTGAAGATCCATA	7326
QY	7021	GGAGATGCTTTCAACAGCTGGCCATTAACTCCTTTGGCCAGCCCCCCCCAAGCGCGAT	7080
Db	7327	GGAGATGCTTTCAACAGCTGGCCATTAAACCTTCGGCCAGATCCCCCAAGCGCGAT	7386
QY	7081	TCAGGCTTTTCAAGGGGGGGGGGGCGTCGGAATTCCGGCAGTCAAGAGCCTCCGATGAG	7140
Db	7387	TCAGGCTTTTCAAGGGGGGGGGGGCGTCGGAATTCGGCAGTCAAGAGCCTCCGATGAG	7446
QY	7141	TTGGCCCTTTTCAAGAGAGGTTCCATCTCTTTCATGCCCCCTCGAGGGGGAGCTTGA	7200
Db	7447	TTGGCTCTTTTCAAGAGAGGTTCCATTCTCCATGCCCCCTTGAAGGGGAGCCTGGG	7506
QY	7201	GATCACAACCTGAGAGCTGAGACAGTATAGACCCCAACCCCCCCCAAGGGGGGGTGGCA	7260
Db	7507	GATCACAACCTGAGAGCCCCGAGACAGTATAGATTCACCTCCCCCAAGGGGGAGACTA	7566
QY	7261	GCTCCGGCTTGAAGCTGGGGTCTCTGATCTACTTGTTCGAGAGAGAGAGACTCCGTCGTG	7320
Db	7567	GCTCCGGCTTGAAGCTGGGGTCTTTGGTCGACTTGTCCGAGGAGATGACTCCGTCGTG	7626
QY	7321	TGCTGCTCATGTATATCTCTTGAACGGGGCTCTAATACTCTTGTAGTCCCGAAGAG	7380
Db	7627	TGCTGCTCATAGTGTACTCTTGAACGGGGCTCTAATACTCTTGTAGTCCCGAAGAA	7686
QY	7381	GAGAAGTTACGATTAACCCCTTGAGCAACTCCCTGTGCGATATCAACAAGAGTTAC	7440
Db	7687	GAGAAGTTACCAATTAACCCCTTGAGCACTCGTGTGCGATATCAACAAGAGTTAC	7746
QY	7441	TGTACACACAAGAGCGGCTCACTAAGGCTTAATAAGGTAACTTTGATAGATGCAA	7500
Db	7747	TGTACTACATCAAGAGAGGCTCACTGAGGGCTTAATAAGGTAACTTTGACAGATGCAA	7806
QY	7501	GTGCTCGACCTCTACTAGACTCACTCACTTAAGGACATTAAGCTAAGGGCTCCCAAGTC	7560
Db	7807	GTGCTCGAGGCCATTAAGACTCACTCTTAAGGACATCAAGCTAAGGGCTCTTAAGGTC	7866
QY	7561	ACCGAAGGCTCTCACCATGAGAGAGGCTTGCCAGTTAACCCCAACCCCAATTCTGACAGA	7620
Db	7867	AGTGAAGGCTCTCACTTAAGAGAGGCTTGCCAGTTGACCCCAACCCCAATTCTGACAGA	7926
QY	7621	TCTAATATTTGGGTTTGGGGCTTAAGAGGTCCGCAAGCTTGTCCGGAGAGGCGCTTAACAC	7680
Db	7927	TCCAAATATTTGGATTTGGGGCTTAAGAGGTCCGCAAGCTTGTCCGGAGAGGCGCTAACCAC	7986
QY	7681	ATCAAGTCCGATGGAAGAGACTCTCTGAGAGACTCAAGAAACACCAATTTCCACAACATT	7740
Db	7987	ATCAAGTCCGATGGAAGAGACTCTCTGAGAGACTCAAGAAACACCAATTTCTACGACATC	8046
QY	7741	ATGGCAAAATGAGGTGTTCTGCTGTGAGCCCAAGGGGGGCAAGAAAGCACTTCG	7800

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 31, 2005, 14:42:35 : Search time 2859 Seconds
(without alignments)
18840.058 Million cell updates/sec

Title: US-09-980-559-1_COPY_341_9439

Perfect score: 9099

Sequence: 1 atgagcacacaatccctaacc.....tttctactcccgctcg 9099

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: genesegn1980s:*
2: genesegn1990s:*
3: genesegn2000s:*
4: genesegn2001as:*
5: genesegn2001bs:*
6: genesegn2002as:*
7: genesegn2002bs:*
8: genesegn2003as:*
9: genesegn2003bs:*
10: genesegn2003cs:*
11: genesegn2003ds:*
12: genesegn2004as:*
13: genesegn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	9099	100.0	9711	4	AAC86937	Aac86937 Nucleotid
2	9099	100.0	9711	5	AAC86644	Aac86644 Nucleotid
3	9095.8	100.0	9711	4	AAE23486	Aae23486 Infectio
4	8494.2	93.4	9589	2	AAQ38218	Aaq38218 NANBH vir
5	7513.4	82.6	9678	6	ABK88904	Abk88904 Human HCV
6	5652.2	62.1	9511	8	AAQ38221	Aaq38221 NANBH vir
7	5399.2	59.3	9611	5	AAC86645	Aac86645 Nucleotid
8	5399.2	59.3	9611	5	AAC86647	Aac86647 Nucleotid
9	5274.4	58.0	9611	5	AAC86646	Aac86646 Nucleotid
10	5274.4	58.0	9611	5	AAC86648	Aac86648 Nucleotid
11	4160.2	45.7	9436	2	AAQ63499	Aaq63499 Blood tra
12	4141.8	45.5	9413	6	AAQ25517	Aad25517 Hepatitis
13	4141.8	45.5	9413	8	AAQ49655	Aad49655 Hepatitis
14	4141.8	45.5	9413	10	ADP88596	Adp88596 Hepatitis
15	4140.2	45.5	9413	8	AAI53723	Aai53723 Hepatitis
16	4138.6	45.5	9413	2	AAQ81559	Aaq81559 Hepatitis
17	4138.6	45.5	9413	2	AAQ03960	Aat03960 Partial H
18	4136	45.5	9599	2	AAQ24833	Aax24833 Infectio
19	4131.2	45.4	9611	13	ADQ54713	Adq54713 Hepatitis
20	4131.2	45.4	12980	6	ABK87286	Abk87286 Hepatitis

21	4130	6	45.4	9413	2	AAQ80498	Aag80498 DNA encod
22	4129	6	45.4	12980	2	AAV59364	Aav59364 Hepatitis
23	4129	6	45.4	12980	8	ACA62469	Aca62469 DNA encod
24	4128	4	45.4	9618	11	ADN33102	Adn33102 Hepatitis
25	4128	4	45.4	9646	2	AAV59361	Aav59361 Hepatitis
26	4128	4	45.4	9646	6	ABK87285	Abk87285 CDNA enco
27	4128	4	45.4	9646	8	ACA62466	Aca62466 HCV H77 c
28	4121	6	45.3	9599	4	AAQ24832	Aax24832 Infectiou
29	4121	6	45.3	9599	4	AAC86938	Aac86938 Nucleotid
30	4121	6	45.3	9599	12	ADQ56743	Adq56743 Hepatitis
31	4121	6	45.3	9599	12	ADQ64255	Adq64255 Hepatitis
32	4121	6	45.3	10803	10	ABX10617	Abx10617 MKO-2 nuc
33	4120	4	45.3	9599	4	AAQ23491	Aaq23491 Infectiou
34	4118	4	45.3	10803	10	ADQ67945	Adq67945 Modified
35	4115	2	45.2	9401	2	AAQ12270	Aat12270 Hepatitis
36	4115	2	45.2	9401	2	AAQ19981	Aat19981 HCV polyp
37	4115	2	45.2	9401	2	AAQ09989	Aav09989 HCV polyp
38	4115	2	45.2	9401	6	AAQ35043	Aad35043 Hepatitis
39	4115	2	45.2	9401	12	ADQ23106	Adi23106 Hepatitis
40	4115	2	45.2	9401	13	ADR29358	Adr29358 Hepatitis
41	4114	4	45.2	9587	13	ADR82189	Adr82189 Hepatitis
42	4112	4	45.2	9385	6	AAQ25518	Aad25518 Hepatitis
43	4112	4	45.2	9400	2	AAQ21744	Aag21744 Compiled
44	4107	6	45.1	9416	2	AAQ59394	Aax59394 Non-A, no
45	4106	4	45.1	9416	2	AAQ20268	Aaq20268 Non-A, no

ALIGNMENTS

RESULT 1						
AC	AAQ86937	standard; DNA; 9711 BP.				
ID	AAQ86937					
AC	AAQ86937					
XX	AAQ86937					
DT	02-APR-2001	(first entry)				
XX	AAQ86937					
DE	Nucleotide sequence of a hepatitis C virus (HCV) clone genotype 2a.					
KW	Chimeric virus; bovine viral diarrhoea virus; BVDV; hepatitis C virus;					
KM	HCV; vaccine; viral inhibitor; antiviral; ss.					
OS	Hepatitis C virus.					
XX	Key	Location/Qualifiers				
FT	CDS	341..9442				
XX	FT	/*tag= a				
XX	PN	WO200075352-A2.				
XX	PD	14-DEC-2000.				
XX	PF	02-JUN-2000; 2000WO-US015527.				
XX	PR	04-JUN-1999; 99US-0137817P.				
XX	PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.				
XX	PI	Nam J, Bukh J, Emerson SU, Purcell RH;				
XX	DR	WPI; 2001-071081/08.				
XX	DR	P-RSDB; AAB1168.				
XX	PT	New nucleic acid comprising a chimeric bovine viral diarrhoea virus genome				
XX	PT	in which the (non-)structural region has been replaced by hepatitis C				
XX	PT	virus (HCV) genome useful for treating or preventing HCV signs and				
XX	PS	symptoms.				
XX	PS	Disclosure; Page 82-85; 97pp; English.				
XX	CC	The specification describes a nucleic acid comprising a chimeric virus				
XX	CC	genome, specifically bovine viral diarrhoea virus (BVDV) genome in which				

the (non-)structural region has been replaced by the (non-)structural region of a hepatitis C virus (HCV) genome. The nucleic acids comprising the chimeric virus and the chimeric virus are useful for identifying cell lines capable of supporting the replication of these chimeric viruses, in screening for neutralizing antibodies to HCV of different genotypes, in the production of HCV-BDV virions, for the development of inactivated or attenuated vaccines to prevent HCV-BDV in a mammal, in studying the molecular properties of HCV indirectly in vitro, and in identifying inhibitors of viral enzyme activity which would be useful as antiviral agents. Formulations or compositions comprising the chimeric viruses may be used to treat or prevent the signs and symptoms of HCV. The present sequence represents a HCV clone, which is used to construct chimeric nucleic acids of the invention

Sequence 9711 BP, 1987 A; 2853 C; 2668 G; 2203 T; 0 U; 0 Other;

Query Match 100.0%; Score 9099; DB 4; Length 9711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9099; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAGCACAATCTCTAACTCAAGAAAAACAAGAAAAACCAACCGTCGCCACA 60
341 ATGAGCACAATCTCTAACTCAAGAAAAACAAGAAAAACAACCGTCGCCACA 400
61 GACGTTAAGTTCCGGGCGCGCGCAATCGTTGGCGGAGTATCTTGTCCGCGAAG 120
401 GACGTTAAGTTCCGGGCGCGCGCGCAATCGTTGGCGGAGTATCTTGTCCGCGAAG 460
121 GGGCCGAGGTTGGGTTGGGCGCGCAAGAAAGCTTGGAGCGGTGCCAGCGTGA 180
461 GGGCCGAGGTTGGGTTGGGCGCGCAAGAAAGCTTGGAGCGGTGCCAGCGTGA 520
181 AGGCGCAGCCCATCTTAAAGATGCGCTCACTGCGCAATCTTGGGAAAAACAAGA 240
521 AGGCGCAGCCCATCTTAAAGATGCGCTCACTGCGCAATCTTGGGAAAAACAAGA 580
241 TACCCCTGAGCCCTATACGGAATAGAGGACTCGGCTGGCGAGATGCTCTCTCCC 300
581 TACCCCTGAGCCCTATACGGAATAGAGGACTCGGCTGGCGAGATGCTCTCTCCC 640
301 CGAGGTTCCGCTCTTGGGGGCGCAATGACCCCGCGCATAGGTCGCGCAAGTGGGT 360
641 CGAGGTTCCGCTCTTGGGGGCGCAATGACCCCGCGCATAGGTCGCGCAAGTGGGT 700
361 AAGGTCATGATACCTTAACGTCGCGCTTTGCGCACTATGAGGTTACATCTCTGTG 420
701 AAGGTCATGATACCTTAACGTCGCGCTTTGCGCACTATGAGGTTACATCTCTGTG 760
421 GGGCGCCGCTCGGCGGCGTCCAGAGCTCTCGCGCATGCGTGAAGTCTTGAAGAC 480
761 GGGCGCCGCTCGGCGGCGTCCAGAGCTCTCGCGCATGCGTGAAGTCTTGAAGAC 820
481 GGGGTTAATTTGCAACAGGGAATTAACCGGTTGCTCTTTCTATCTTCTGTGGCC 540
821 GGGGTTAATTTGCAACAGGGAATTAACCGGTTGCTCTTTCTATCTTCTGTGGCC 880
541 CTGCTGTCTGTGATACCAACCCCGCTCTCGCTGCGAAGTGAAGAACATGACCGGC 600
881 CTGCTGTCTGTGATACCAACCCCGCTCTCGCTGCGAAGTGAAGAACATGACCGGC 940
601 TACATGTGATCTAAGACTGACCAATGACATGACATTAACCTGCGAGCTCTGTCTC 660
941 TACATGTGATCTAAGACTGACCAATGACATGACATTAACCTGCGAGCTCTGTCTC 1000
661 CTGCAAGTCCCGGGTGGTCCCGTGGCAAGAGTGGGAATGATCTCAGTGTGATA 720
1001 CTGCAAGTCCCGGGTGGTCCCGTGGCAAGAGTGGGAATGATCTCAGTGTGATA 1060
721 CCGGCTCTACCGAATGTGGCGTGGAGCGGCCCGGCTCTACGCAAGGCTTGGCGAG 780
1061 CCGGCTCTACCGAATGTGGCGTGGAGCGGCCCGGCTCTACGCAAGGCTTGGCGAG 1120
781 CACATGACATGTTGTGATGTCGCCACAGCTTGTCTGCTTACGTTGGGAGCCTC 840

1121 CACATGACATGTTGTGATGTCGCCACAGCTTGTCTGCGCTTACGTTGGGAGCCTC 1180
841 TGGGTTGGGATGATGCTCGACAGCCCAATGTTTATTTGTCGCGGACGACCACTGTT 900
1181 TGGGTTGGGATGATGCTCGACAGCCCAATGTTTATTTGTCGCGGACGACCACTGTT 1240
901 GTCGAAGCTGCAATTTGCTCATCTACCTGTGACCATGACATGACACCGATGATG 960
1241 GTCGAAGCTGCAATTTGCTCATCTACCTGTGACCATGACATGACACCGATGATG 1300
961 GACATGATGATACTGTGTGCGCCAGGCTACCATATCTTGGCGTACCGATGCGTTC 1020
1301 GACATGATGATACTGTGTGCGCCAGGCTACCATATCTTGGCGTACCGATGCGTTC 1360
1021 CCGGAGTCAATTAAGACATCATTAACGAGGCTCATTTGGGAGCGTACGTTGCGCT 1080
1361 CCGGAGTCAATTAAGACATCATTAACGAGGCTCATTTGGGAGCGTACGTTGCGCT 1420
1081 TACTTCTTAAGAGGAGCGTGGCGAAAGTGTGATCTTCTGTGGCGCGCGAG 1140
1421 TACTTCTTAAGAGGAGCGTGGCGAAAGTGTGATCTTCTGTGGCGCGCGAG 1480
1141 GTGAGCGCGGACCCATTACTGTGGGAGTGTGCGCGGACGACACCGGCGCTCAC 1200
1481 GTGAGCGCGGACCCATTACTGTGGGAGTGTGCGCGGACGACACCGGCGCTCAC 1540
1201 AGCTATTGACATGAGGCGCCCAAGGCAAAATCAGCTGTTTACACCAATGAGCAG 1260
1541 AGCTATTGACATGAGGCGCCCAAGGCAAAATCAGCTGTTTACACCAATGAGCAG 1600
1261 CACATGAACCGACCGCCCTGAACTGCAATGACTCTTGCACACCGGCTTTATCGGCT 1320
1601 CACATGAACCGACCGCCCTGAACTGCAATGACTCTTGCACACCGGCTTTATCGGCT 1660
1321 CTGTTTCAACCAACGCTTCAACTGCTGAGATGTCGCCGAACGATGTCGCCGCTCC 1380
1661 CTGTTTCAACCAACGCTTCAACTGCTGAGATGTCGCCGAACGATGTCGCCGCTCC 1720
1381 AGTATGAGGCTTCCGGGTGGGATGAGGCGCTTGCATATAGAGTATGACCAAT 1440
1721 AGTATGAGGCTTCCGGGTGGGATGAGGCGCTTGCATATAGAGTATGACCAAT 1780
1441 CCAAGGATATGAGACCTTATTTGCTGCACTACCAACCAAGCATGTCGCTGCTCC 1500
1781 CCAAGGATATGAGACCTTATTTGCTGCACTACCAACCAAGCATGTCGCTGCTCC 1840
1501 GCGAAGATGTCGTCGCGGCGGAGTACTGTTCAACCCGACGCGCATGTCGTCGAG 1560
1841 GCGAAGATGTCGTCGCGGCGGAGTACTGTTCAACCCGACGCGCATGTCGTCGAG 1900
1561 ACCGAGGCTTGAAGCGCCCACTTACAGTGGGAGGAGATGAGACAGATGCTTCTTA 1620
1901 ACCGAGGCTTGAAGCGCCCACTTACAGTGGGAGGAGATGAGACAGATGCTTCTTA 1960
1621 TTGAAGACACTGACACCGCTGAGGATGATGTTGCGCTGACGTCGATGAATCTTCT 1680
1961 TTGAAGACACTGACACCGCTGAGGATGATGTTGCGCTGACGTCGATGAATCTTCT 2020
1681 GGTACACCAAGCTTGGCGGCGACCAACCTGCGGATGATGAGCTGATCAACGCGCAG 1740
2021 GGTACACCAAGCTTGGCGGCGACCAACCTGCGGATGATGAGCTGATCAACGCGCAG 2080
1741 ACCGAGCTGTTGCGCCACGAGCTGTTTGAAGACATCTGATACCACTTACCTCAAA 1800
2081 ACCGAGCTGTTGCGCCACGAGCTGTTTGAAGACATCTGATACCACTTACCTCAAA 2140
1801 TGGGCTCTGAGGCTTGGCTCAAGCGCAAGTGTCTGATGACATCCCTTACAGGCTCT 1860
2141 TGGGCTCTGAGGCTTGGCTCAAGCGCAAGTGTCTGATGACATCCCTTACAGGCTCT 2200
1861 CATTAACCTTGCACATGATTAATTAATCTTCAAAATAGATGATGAGGAGGCTT 1920

D 2201 CATTACCCCTGCACAGTAACTATACCATCTTCAAAATAAGATGTATGTGGAGGGGTT 2260
Q 1921 GAGACAGGCTCAAGGCTGCATGCAATTTCACTGTGGGGATCGTTGCACTTGGAGAC 1980
D 2261 GAGACAGGCTCAAGGCTGCATGCAATTTCACTGTGGGGATCGTTGCACTTGGAGAC 2320
Q 1981 AGAGACAGAAGTCAAGTCTCTTTGTGCACTCCACACGGAATGGGCAATTTACT 2040
D 2321 AGAGACAGAAGTCAAGTCTCTTTGTGCACTCCACACGGAATGGGCAATTTACT 2380
Q 2041 TGCTTTACTGCGACCTGCGCGCTTGCGACTGCTCTTCCACTCCACAAATATC 2100
D 2381 TGCTTTACTGCGACCTGCGCGCTTGCGACTGCTCTTCCACTCCACAAATATC 2440
Q 2101 GTGACGCTACAAATTCATGTATGCGCTATCACTGCGCTCAAAATATCATGTCGATGG 2160
D 2441 GTGACGCTACAAATTCATGTATGCGCTATCACTGCGCTCAAAATATCATGTCGATGG 2500
Q 2161 GAGTGGTAAATTAATCTTATTCTGCTCTTAGCGGACCGAGGGTTGCGCTGTATGG 2220
D 2501 GAGTGGTAAATTAATCTTATTCTGCTCTTAGCGGACCGAGGGTTGCGCTGTATGG 2560
Q 2221 ATGCTCATCTTGTGGGCCAGGCCGAAAGCAGCACTAGAGAAGCTGTCATCTTGCACGCT 2280
D 2561 ATGCTCATCTTGTGGGCCAGGCCGAAAGCAGCACTAGAGAAGCTGTCATCTTGCACGCT 2620
Q 2281 GCCAGCGCAGCTAGCTGCAATGGTCTCTATATTTTGTCACTTTTTCGTGGCTGTGG 2340
D 2621 GCCAGCGCAGCTAGCTGCAATGGTCTCTATATTTTGTCACTTTTTCGTGGCTGTGG 2680
Q 2341 TACATCAAGGGGTGGGGATAGTCCCTTAGCTACCTATTCCTCACTGGCGTGGTCCCTT 2400
D 2681 TACATCAAGGGGTGGGGATAGTCCCTTAGCTACCTATTCCTCACTGGCGTGGTCCCTT 2760
Q 2401 AGCCTACTGCTCTTAGCAATTCGCCCAACAGGCTTATGTGAAGCATGTGTCATGGC 2460
D 2741 AGCCTACTGCTCTTAGCAATTCGCCCAACAGGCTTATGTGAAGCATGTGTCATGGC 2800
Q 2461 CAGATAGGAGCGGCTGCTGTGTAATGATCACTCTTTACTCTACCCCGGGATATAG 2520
D 2801 CAGATAGGAGCGGCTGCTGTGTAATGATCACTCTTTACTCTACCCCGGGATATAG 2860
Q 2521 ACCCTTCTCAGCGGGTTTTGTGGTGTGCTATCTTCAAGCCCTGGGGGAAGCTATG 2580
D 2861 ACCCTTCTCAGCGGGTTTTGTGGTGTGCTATCTTCAAGCCCTGGGGGAAGCTATG 2920
Q 2581 GTCAGAGAGTGGGCAACCATATGCAAGTGGCGCGGTGATGGCATATATGGGCC 2640
D 2921 GTCAGAGAGTGGGCAACCATATGCAAGTGGCGCGGTGATGGCATATATGGGCC 2980
Q 2641 GTGGCATATTTCTAACCAAGGTGTGTTGACATAACCAAGTGGCTTTGGCGGTCTT 2700
D 2981 GTGGCATATTTCTAACCAAGGTGTGTTGACATAACCAAGTGGCTTTGGCGGTCTT 3040
Q 2701 GGGGCTGCTTAACCTCTAAAGAGTCTTGAACGCGGCTGCTCTTCAAGGGCTCAC 2760
D 3041 GGGGCTGCTTAACCTCTAAAGAGTCTTGAACGCGGCTGCTCTTCAAGGGCTCAC 3100
Q 2761 GCTCTACTGAGATGTGCAACCATATGGCAAGGATCTCGCGGGGCAAGTACGTCAGATG 2820
D 3101 GCTCTACTGAGATGTGCAACCATATGGCAAGGATCTCGCGGGGCAAGTACGTCAGATG 3160
Q 2821 GCGCTACTAGCCCTTGGCAGGTGGACTGGGCACTTATACCTATATGACCACTCACCCCTATG 3220
D 3161 GCGCTACTAGCCCTTGGCAGGTGGACTGGGCACTTATACCTATATGACCACTCACCCCTATG 3280
Q 2881 TCGGATTTGGGCTGTAGTGGCTGCGGGAACCTGGCGGTGCGCTTGAAGCCATATCTTC 2940
D 3221 TCGGATTTGGGCTGTAGTGGCTGCGGGAACCTGGCGGTGCGCTTGAAGCCATATCTTC 3280
Q 2941 AGTCCGATGAGAAAGATCATTTGTCTGGGAGCGGAGACAGCTGCTTGTGGGACATTT 3000
D 3281 AGTCCGATGAGAAAGATCATTTGTCTGGGAGCGGAGACAGCTGCTTGTGGGACATTT 3340

Q 3001 TTAACGGAATTCGCCGTGCGGCCGACTTGTGTCGGGAGGTCCTCTTGGCCGAGCTGAT 3060
D 3341 TTAACGGAATTCGCCGTGCGGCCGACTTGTGTCGGGAGGTCCTCTTGGCCGAGCTGAT 3400
Q 3061 GGCCTATACCTCAAGGGGTGAGTCTTCTGCCCCATCACTGCTTACGCCACAGACACA 3120
D 3401 GGCCTATACCTCAAGGGGTGAGTCTTCTGCCCCATCACTGCTTACGCCACAGACACA 3460
Q 3121 CGTGGCTTTTGGGACATATGTGTGAGATGACGCGGGGCGGACAAAGACAGAGCT 3180
D 3461 CGTGGCTTTTGGGACATATGTGTGAGATGACGCGGGGCGGACAAAGACAGAGCT 3520
Q 3181 GGGGAAATTCAGGTCCTGTGCAAGTCACTCAAGCTCTTCCGGAATCATCATCTCGGGG 3240
D 3521 GGGGAAATTCAGGTCCTGTGCAAGTCACTCAAGCTCTTCCGGAATCATCATCTCGGGG 3580
Q 3241 GTTTTGTGACTGTCTACATGAGAGCTGGCAACAAAGCTTGGCGGCTCAAGGGGTCCG 3300
D 3581 GTTTTGTGACTGTCTACATGAGAGCTGGCAACAAAGCTTGGCGGCTCAAGGGGTCCG 3640
Q 3301 GTCAAGCAGATGTACTCAAGTGTGAGGGGGACTTATAGGTGGCCAGCCCTCTGGG 3360
D 3641 GTCAAGCAGATGTACTCAAGTGTGAGGGGGACTTATAGGTGGCCAGCCCTCTGGG 3700
Q 3361 ACTAAATCTTTGGAGCGGTGCAAGTGTGAGCGGTGCACTGTACCTGTACGCGGGAAC 3420
D 3701 ACTAAATCTTTGGAGCGGTGCAAGTGTGAGCGGTGCACTGTACCTGTGTACCGGGAAC 3760
Q 3421 GCTGATGTCACTCCGCTCGAAGACGCGGGGACAAAGCGGGAGGCTACTTCCCGAGA 3480
D 3761 GCTGATGTCACTCCGCTCGAAGACGCGGGGACAAAGCGGGAGGCTACTTCCCGAGA 3820
Q 3481 CTTCTTTCACTTTGAAGGGGTCTCTAGAGGCCCGGTGTATGCCCCAGGGGCAAGCT 3540
D 3821 CTTCTTTCACTTTGAAGGGGTCTCTAGAGGCCCGGTGTATGCCCCAGGGGCAAGCT 3880
Q 3541 GTGGAGTCTTCCGGGAGCTGTGTGCTCTCGGGGCGTGGCTAAGTGCATATGATTTATC 3600
D 3881 GTGGAGTCTTCCGGGAGCTGTGTGCTCTCGGGGCGTGGCTAAGTGCATATGATTTATC 3940
Q 3601 CCGGTTGAGACACTCGACATGTCACGCGGTGCCCACTTTATGTGACAAAGACACACA 3660
D 3941 CCGGTTGAGACACTCGACATGTCACGCGGTGCCCACTTTATGTGACAAAGACACACA 4000
Q 3661 CTTGCTGTGCCCAAGACTTATCAGGTGGGTACTTGCATGCCGACTGGCAGTGGAAAG 3720
D 4001 CTTGCTGTGCCCAAGACTTATCAGGTGGGTACTTGCATGCCGACTGGCAGTGGAAAG 4060
Q 3721 AGCAACCAAGTTCCTGTGCGATATGCTGCTCAGGGGTATTAAGTGTAGTGAATCCC 3780
D 4061 AGCAACCAAGTTCCTGTGCGATATGCTGCTCAGGGGTATTAAGTGTAGTGAATCCC 4120
Q 3781 TCAATGCTGCTCAACCTGCGGGTCTTGAAGGCACTATGATATCC 3840
D 4121 TCAATGCTGCTCAACCTGCGGGTCTTGAAGGCACTATGATATCC 4180
Q 3841 AACATTAAGACTGAGTCAAGACTGTGACGACCGGGGCGCCATCAGTACTCCACATAT 3900
D 4181 AACATTAAGACTGAGTCAAGACTGTGACGACCGGGGCGCCATCAGTACTCCACATAT 4240
Q 3901 GGGAAATTTCTCCCGCATGGGGGCTGTGCGGGGCGGCTTACGACATCATATATGAT 3960
D 4241 GGGAAATTTCTCCCGCATGGGGGCTGTGCGGGGCGGCTTACGACATCATATATGAT 4300
Q 3961 GAATGCCATGCGGTGACTTACCAACCATCTTGGATGGGAAGTCTTTGATCAAGCA 4020
D 4301 GAATGCCATGCGGTGACTTACCAACCATCTTGGATGGGAAGTCTTTGATCAAGCA 4360
Q 4021 GAGACAGCTGGGTCAAGCTTAACTGTGTGCTGACGCTACGCTTGGGTCAGTGACA 4080
D 4361 GAGACAGCTGGGTCAAGCTTAACTGTGTGCTGACGCTTGGGTCAGTGACA 4420

4081 ACCCCCAACCCCAATAGAGAGGTGGCCCTTGGGAGAGGGAGATCCCTTCTAT 4140
Db ACCCCCAACCCCAATAGAGAGGTGGCCCTTGGGAGAGGGAGATCCCTTCTAT 4480
4141 GGGAGGGGATTTCCCTGTCTTACATCAAGGAGGAAAGACATCTGATCTTCTGGCATTTCA 4200
Db GGGAGGGGATTTCCCTGTCTTACATCAAGGAGGAAAGACATCTGATCTTCTGGCATTTCA 4540
4201 AAGAAAAAGTGTGACGAGCTCGCGCGCGCCCTTGGGGGTATGGCTTGAATCAGTGGCA 4260
Db AAGAAAAAGTGTGACGAGCTCGCGCGCGCCCTTGGGGGTATGGCTTGAATCAGTGGCA 4600
4261 TACTACAGAGGAGTGGAGCGTCTCCGTATACCAACTCAGGAGAGAGTATGATGCTGCGCC 4660
4321 ACCGACGCGCTCATGACAGGGATATCTGGGAGCTTTGACTCGGATCTGATGATGCAAGCTA 4380
Db ACCGACGCGCTCATGACAGGGATATCTGGGAGCTTTGACTCGGATCTGATGATGCAAGCTA 4720
4381 GCGGTCACTCAAGTTGTAGACTTCAGTTTAAAGCCACATTCACCATTAACACAGATT 4440
Db GCGGTCACTCAAGTTGTAGACTTCAGTTTAAAGCCACATTCACCATTAACACAGATT 4780
4721 GCGGTCACTCAAGTTGTAGACTTCAGTTTAAAGCCACATTCACCATTAACACAGATT 4780
4441 GTCCCTCAAGACGCTGTCTCACTGACGACGCGCGGGGTGCGACGGGTAGGGGAAAGCTG 4500
Db GTCCCTCAAGACGCTGTCTCACTGACGACGCGCGGGGTGCGACGGGTAGGGGAAAGCTG 4840
4501 GGCATTATAGATATGTTTCACTGGTATGAGAGGAGCTTCAAGGAGTGTGACAGTGTAGT 4560
Db GGCATTATAGATATGTTTCACTGGTATGAGAGGAGCTTCAAGGAGTGTGACAGTGTAGT 4900
4561 CTCTGTGAGTGTACGACGAGGAGCGGAGCTGATGAGCTCAGACATTCGAGACACC 4620
Db CTCTGTGAGTGTACGACGAGGAGCGGAGCTGATGAGCTCAGACATTCGAGACACC 4960
4621 GTCAGGCTCAGGGCGTATTTCAACACGCGCGGTTTGGCTGTGTGCTCAAGACCATCTTGAG 4680
4961 GTCAGGCTCAGGGCGTATTTCAACACGCGCGGTTTGGCTGTGTGCTCAAGACCATCTTGAG 5020
4681 TTTTGGGAGGAGTTTTCACCGGCTCAGACATGATGACCTCTTCTTCCCAACA 4740
Db TTTTGGGAGGAGTTTTCACCGGCTCAGACATGATGACCTCTTCTTCCCAACA 5080
5021 TTTTGGGAGGAGTTTTCACCGGCTCAGACATGATGACCTCTTCTTCCCAACA 5080
4741 AAGCAATCGGGGAAAAATTTGSCATCTTAAACAGCTACAGGCTACAGTGTGCTAGG 4800
Db AAGCAATCGGGGAAAAATTTGSCATCTTAAACAGCTACAGGCTACAGTGTGCTAGG 5140
5081 AAGCAATCGGGGAAAAATTTGSCATCTTAAACAGCTACAGGCTACAGTGTGCTAGG 5140
4801 GCCAAAGCCCCCCCCGCTCTGGGAGCTCATGTGGAAGTGTGACTCGACTAAGCCC 4860
Db GCCAAAGCCCCCCCCGCTCTGGGAGCTCATGTGGAAGTGTGACTCGACTAAGCCC 5200
5141 GCCAAAGCCCCCCCCGCTCTGGGAGCTCATGTGGAAGTGTGACTCGACTAAGCCC 5200
4861 AACCTCGTGGGCCCCACCTCTCTGTACCGGCTTGGGCTCTGTTTCCAAAGAGGTACCC 4920
Db AACCTCGTGGGCCCCACCTCTCTGTACCGGCTTGGGCTCTGTTTCCAAAGAGGTACCC 5260
5201 AACCTCGTGGGCCCCACCTCTCTGTACCGGCTTGGGCTCTGTTTCCAAAGAGGTACCC 5260
4921 CTCACACATCCCGTGTGACGAATATACCTGCACTGTGATGCAAGCGACCTTGAAGTATG 4980
Db CTCACACATCCCGTGTGACGAATATACCTGCACTGTGATGCAAGCGACCTTGAAGTATG 5320
5261 CTCACACATCCCGTGTGACGAATATACCTGCACTGTGATGCAAGCGACCTTGAAGTATG 5320
4981 ACCAGCACAATGGGTCTTGGCAGGGGGAATCTTGGCGCGCTGCGCGGTATTTGCTTGGCG 5040
Db ACCAGCACAATGGGTCTTGGCAGGGGGAATCTTGGCGCGCTGCGCGGTATTTGCTTGGCG 5380
5321 ACCAGCACAATGGGTCTTGGCAGGGGGAATCTTGGCGCGCTGCGCGGTATTTGCTTGGCG 5380
5041 ACCGGGTGTGTTTGATCATCGGCGCTTGCACATTTAACAGCGAGCGGTGTGGCGCG 5100
Db ACCGGGTGTGTTTGATCATCGGCGCTTGCACATTTAACAGCGAGCGGTGTGGCGCG 5440
5381 ACCGGGTGTGTTTGATCATCGGCGCTTGCACATTTAACAGCGAGCGGTGTGGCGCG 5440
5101 GACAAGAGGTCTCTATAGAGCTTTTGTATGATGAGTGAAGTGTGCTCTTGAAGGCGCT 5160
Db GACAAGAGGTCTCTATAGAGCTTTTGTATGATGAGTGAAGTGTGCTCTTGAAGGCGCT 5500
5441 GACAAGAGGTCTCTATAGAGCTTTTGTATGATGAGTGAAGTGTGCTCTTGAAGGCGCT 5500
5161 CTGATTGAAGAGGGGAGCGGATGAGCGAGATGCTGAAGTCCAAAGCTTATTG 5220

5501 CTGATTGAAGAGGGGAGCGGATAGCGGAGTGTGAATGCCAAGATCCAAAGCTTATTG 5560
Db CTGATTGAAGAGGGGAGCGGATAGCGGAGTGTGAATGCCAAGATCCAAAGCTTATTG 5560
5501 CTGATTGAAGAGGGGAGCGGATAGCGGAGTGTGAATGCCAAGATCCAAAGCTTATTG 5560
5221 CAGCAAGCTTCCAAAGAGCTCAAGCATTAACAACCACTGTGAGGCTTATGGCCCAAG 5280
Db CAGCAAGCTTCCAAAGAGCTCAAGCATTAACAACCACTGTGAGGCTTATGGCCCAAG 5620
5561 CAGCAAGCTTCCAAAGAGCTCAAGCATTAACAACCACTGTGAGGCTTATGGCCCAAG 5620
5281 GTAGAACAAATTCTGGGCGCAACATGTGGAATCTTATTAAGCGGATCCAAATCTCGCA 5340
Db GTAGAACAAATTCTGGGCGCAACATGTGGAATCTTATTAAGCGGATCCAAATCTCGCA 5680
5621 GTAGAACAAATTCTGGGCGCAACATGTGGAATCTTATTAAGCGGATCCAAATCTCGCA 5680
5341 GGAATATCAACATGCGCAGGAAACCTGCAATAGCTTCCATGATGGCGTTCAATGCGCGC 5400
Db GGAATATCAACATGCGCAGGAAACCTGCAATAGCTTCCATGATGGCGTTCAATGCGCGC 5740
5681 GGAATATCAACATGCGCAGGAAACCTGCAATAGCTTCCATGATGGCGTTCAATGCGCGC 5740
5401 CTCACAGTCCGCTGTCAACAGACACATATCTTCTCAACATTTTGGGGGCTGGCTA 5460
Db CTCACAGTCCGCTGTCAACAGACACATATCTTCTCAACATTTTGGGGGCTGGCTA 5800
5741 CTCACAGTCCGCTGTCAACAGACACATATCTTCTCAACATTTTGGGGGCTGGCTA 5800
5461 GCATCCCAAAATTTGACCAACCCCGGAGGAGCACTGGCTGTGTCAATGAGCTTATGAGGA 5520
Db GCATCCCAAAATTTGACCAACCCCGGAGGAGCACTGGCTGTGTCAATGAGCTTATGAGGA 5860
5801 GCATCCCAAAATTTGACCAACCCCGGAGGAGCACTGGCTGTGTCAATGAGCTTATGAGGA 5860
5521 GCTGCGGTAGGCAATATAGGCTTATGATAGTGTGTGATGATCTCTGGCAGGCTATGGT 5580
Db GCTGCGGTAGGCAATATAGGCTTATGATAGTGTGTGATGATCTCTGGCAGGCTATGGT 5920
5861 GCTGCGGTAGGCAATATAGGCTTATGATAGTGTGTGATGATCTCTGGCAGGCTATGGT 5920
5581 GCGGGGATTTGAGGGGCTCTCGTGCATTCAGATATGATCTGGCAGAGAGCCCTCCATG 5640
Db GCGGGGATTTGAGGGGCTCTCGTGCATTCAGATATGATCTGGCAGAGAGCCCTCCATG 5980
5921 GCGGGGATTTGAGGGGCTCTCGTGCATTCAGATATGATCTGGCAGAGAGCCCTCCATG 5980
5641 GAGGATGTGCTCAACTTGTGCTGCTGGAATTTGTCTCCGAGTGCCTTGTGATGAGGATG 5700
Db GAGGATGTGCTCAACTTGTGCTGCTGGAATTTGTCTCCGAGTGCCTTGTGATGAGGATG 6040
5981 GAGGATGTGCTCAACTTGTGCTGCTGGAATTTGTCTCCGAGTGCCTTGTGATGAGGATG 6040
5701 ATCTGGCGGAGCTTCTGGGCGCAACGTTGGGACCGGGGGAAGGCCCTGCCAATGGATG 5760
Db ATCTGGCGGAGCTTCTGGGCGCAACGTTGGGACCGGGGGAAGGCCCTGCCAATGGATG 6100
6041 ATCTGGCGGAGCTTCTGGGCGCAACGTTGGGACCGGGGGAAGGCCCTGCCAATGGATG 6100
5761 AATAGACTCATGCTCTTGGCTTCCAGAGGAAATCAAGCTGAGCCGACCCCACTAGTACG 5820
Db AATAGACTCATGCTCTTGGCTTCCAGAGGAAATCAAGCTGAGCCGACCCCACTAGTACG 6160
6101 AATAGACTCATGCTCTTGGCTTCCAGAGGAAATCAAGCTGAGCCGACCCCACTAGTACG 6160
5821 GAGTGGATGCTGTGAGAGCTGTGACCAACTACTTGTGCTCTGTGCAAAAGGGTACAG 5880
Db GAGTGGATGCTGTGAGAGCTGTGACCAACTACTTGTGCTCTGTGCAAAAGGGTACAG 6220
6161 GAGTGGATGCTGTGAGAGCTGTGACCAACTACTTGTGCTCTGTGCAAAAGGGTACAG 6220
5881 CTCAAGAACTCCAACTGGAATTAAGAGATGAGGATGAGGATGAGGATGAGGATGAGGATG 5940
Db CTCAAGAACTCCAACTGGAATTAAGAGATGAGGATGAGGATGAGGATGAGGATGAGGATG 6280
6221 CTCAAGAACTCCAACTGGAATTAAGAGATGAGGATGAGGATGAGGATGAGGATGAGGATG 6280
5941 CTCGCGATGTGTGGGACTGGGTTTGCAACATCTTAACAGACTTTTAAAAATTTGGCTGACC 6000
Db CTCGCGATGTGTGGGACTGGGTTTGCAACATCTTAACAGACTTTTAAAAATTTGGCTGACC 6340
6281 CTCGCGATGTGTGGGACTGGGTTTGCAACATCTTAACAGACTTTTAAAAATTTGGCTGACC 6340
6001 TCCAAATTAATCCCAAGATGCGCGGCTTCCCTTGTCTCTGTGCAAAAGGGTACAG 6060
Db TCCAAATTAATCCCAAGATGCGCGGCTTCCCTTGTCTCTGTGCAAAAGGGTACAG 6400
6341 TCCAAATTAATCCCAAGATGCGCGGCTTCCCTTGTCTCTGTGCAAAAGGGTACAG 6400
6061 GGGGTGTGGGCGGCACTGGCATCATGACCAACGAGTGTCTTGGCGCGCAATATCTCT 6120
Db GGGGTGTGGGCGGCACTGGCATCATGACCAACGAGTGTCTTGGCGCGCAATATCTCT 6460
6401 GGGGTGTGGGCGGCACTGGCATCATGACCAACGAGTGTCTTGGCGCGCAATATCTCT 6460
6121 GGCATGTCCGCTTGGGCTCATGAGAAATCAAGGAGCTTAAGACTGATGATATCTGG 6180
Db GGCATGTCCGCTTGGGCTCATGAGAAATCAAGGAGCTTAAGACTGATGATATCTGG 6520
6461 GGCATGTCCGCTTGGGCTCATGAGAAATCAAGGAGCTTAAGACTGATGATATCTGG 6520
6181 CAGGGGACCTTCTCTATCAATTTGTTACAGGAGGAGCAAGTGTGCTGCGGAAACCCGCGCA 6240
Db CAGGGGACCTTCTCTATCAATTTGTTACAGGAGGAGCAAGTGTGCTGCGGAAACCCGCGCA 6580
6521 CAGGGGACCTTCTCTATCAATTTGTTACAGGAGGAGCAAGTGTGCTGCGGAAACCCGCGCA 6580
6241 AACTTTAAGTGGCCATCTGAGAGGTGGGCTCAGAGTACCGGAGAGTGAAGCGAGC 6300

Dh 6581 AACTTTAAGTGCCATCTGGAGGGGTGGCGCTCAGAGTACGCGAGGTGACGACGAC 6640
Qy 6301 GGGTCATACCACTACATTAACAGGACTCAGCAGCTGATTAAGTAAAGTCCCTGCAACTA 6360
Dh 6641 GGGTCATACCACTACATTAACAGGACTCAGCAGCTGATTAAGTAAAGTCCCTGCAACTA 6700
Qy 6361 CCTCTCCCGAGTTCTTTCTCTGGGTGAGCGAGTGCAGATCATAAGTTTGGCCCCA 6420
Dh 6701 CCTCTCCCGAGTTCTTTCTCTGGGTGAGCGAGTGCAGATCATAAGTTTGGCCCCA 6760
Qy 6421 CCGAAGCGTTTTCGCGAGTGAAGTCTCGTTCTGCGTTGGGCTTAAATCATTTGTGCTC 6480
Dh 6761 CCGAAGCGTTTTCGCGAGTGAAGTCTCGTTCTGCGTTGGGCTTAAATCATTTGTGCTC 6820
Qy 6481 GGGTCACAGCTTCTTTCGAGCCCTGAAACCGCAGACAGCTAATGATGCTCACTAACA 6540
Dh 6821 GGGTCACAGCTTCTTTCGAGCCCTGAAACCGCAGACAGCTAATGATGCTCACTAACA 6880
Qy 6541 GATTCATCTCATATCAAGCGAGAGATGAGCGCGGCTTAAAGCGGGGTCAACCCCA 6600
Dh 6881 GATTCATCTCATATCAAGCGAGAGATGAGCGCGGCTTAAAGCGGGGTCAACCCCA 6940
Qy 6601 TCCGAGCGAAGCTCTCTGCGAGAGCAGCTATCGGACCATCGCTGCAAGCCACTGACC 6660
Dh 6941 TCCGAGCGAAGCTCTCTGCGAGAGCAGCTATCGGACCATCGCTGCAAGCCACTGACC 7000
Qy 6661 ACCCAGCGCAAGCGCTATGATGTGAGATGAGTGGATGCTAACTGTTCAATGGGGGGGAT 6720
Dh 7001 ACCCAGCGCAAGCGCTATGATGTGAGATGAGTGGATGCTAACTGTTCAATGGGGGGGAT 7060
Qy 6721 GTGATCGGATAGAGTCTGGGTCCAAAGTGTCTTTCGAGCTCTCTGAGCCCAATGCTC 6780
Dh 7061 GTGATCGGATAGAGTCTGGGTCCAAAGTGTCTTTCGAGCTCTCTGAGCCCAATGCTC 7120
Qy 6781 GAAGAAAGAGCGCAGCTTGAAGCTTGCATACATCAGATATCATGCTTCCCAAGAGAGG 6840
Dh 7121 GAAGAAAGAGCGCAGCTTGAAGCTTGCATACATCAGATATCATGCTTCCCAAGAGAGG 7180
Qy 6841 TTTCCACACACTTTTACCGGCTCTGGGCAAGCGCTGATTAACCAACCGCTTTGTGAATCG 6900
Dh 7181 TTTCCACACACTTTTACCGGCTCTGGGCAAGCGCTGATTAACCAACCGCTTTGTGAATCG 7240
Qy 6901 TGGAAAGAGCGCAGATTAACCAACCGGCTTGAAGCTTTCGAGCTCTCTCTGAGG 6960
Dh 7241 TGGAAAGAGCGCAGATTAACCAACCGGCTTGAAGCTTTCGAGCTCTCTCTGAGG 7300
Qy 6961 AAAACCCCGAGCGCTCTCCCAAGAGAGCGCCGAGCAGTGGGCTTAAGTGAAGACTCATA 7020
Dh 7301 AAAACCCCGAGCGCTCTCCCAAGAGAGCGCCGAGCAGTGGGCTTAAGTGAAGACTCATA 7360
Qy 7021 GAGATGCTCTTCAACAGCTGAGCTTAAGTCTTTTGGCCAGCCCCCCCCCAAGGGGAGT 7080
Dh 7361 GAGATGCTCTTCAACAGCTGAGCTTAAGTCTTTTGGCCAGCCCCCCCCCAAGGGGAGT 7420
Qy 7081 TCAAGCTTTTCAAGCGGGGCGGAGCTGCGATTCGCGAGTCAAGCGCTCTGATGAG 7140
Dh 7421 TCAAGCTTTTCAAGCGGGGCGGAGCTGCGATTCGCGAGTCAAGCGCTCTGATGAG 7480
Qy 7141 TTGGCTTTTCGAGAGCAGGTTTCATCTTTCATGCCCCCTCGAGGGGAGACTTGA 7200
Dh 7481 TTGGCTTTTCGAGAGCAGGTTTCATCTTTCATGCCCCCTCGAGGGGAGACTTGA 7540
Qy 7201 GATTCAGACTGAGAGCTGAGAGAGTGAAGCCCAACCCCGCGAGGGGGGGTGA 7260
Dh 7541 GATTCAGACTGAGAGCTGAGAGAGTGAAGCCCAACCCCGCGAGGGGGGGTGA 7600
Qy 7261 GCTCCCGGCTCGAGCTCGGGGTCTTGTCTTACTTGTCTCGAGAGAGAGCAGCTCGTGTG 7320
Dh 7601 GCTCCCGGCTCGAGCTCGGGGTCTTGTCTTACTTGTCTCGAGAGAGAGCAGCTCGTGTG 7660
Qy 7321 TGTGCTCATGTCTATCTCTTGAAGCGGGCTTAATACTCTTGTAGTCCGAAGAG 7380
Dh 7661 TGTGCTCATGTCTATCTCTTGAAGCGGGCTTAATACTCTTGTAGTCCGAAGAG 7720

Qy 7381 GAGAGTTACCGATTAACCCCTTGAGCACTCCGTGTGCGATATCAACAAGGTGTAC 7440
Dh 7721 GAGAGTTACCGATTAACCCCTTGAGCACTCCGTGTGCGATATCAACAAGGTGTAC 7780
Qy 7441 TGTACCAACAAGAGCGCTCACTAAGGCTTAAGAGTTAACTTTTGAATGAGTCAA 7500
Dh 7781 TGTACCAACAAGAGCGCTCACTAAGGCTTAAGAGTTAACTTTTGAATGAGTCAA 7840
Qy 7501 GTGTCGACTCTTACTACAGTCACTGCTTAAAGACATTAAAGTACGGGCTCCAAAGTC 7560
Dh 7841 GTGTCGACTCTTACTACAGTCACTGCTTAAAGACATTAAAGTACGGGCTCCAAAGTC 7900
Qy 7561 ACCGCAAGGCTCCCACTAGAGAGAGGCTTGCAGTTAAACCCCACTTCGACAGA 7620
Dh 7901 ACCGCAAGGCTCCCACTAGAGAGAGGCTTGCAGTTAAACCCCACTTCGACAGA 7960
Qy 7621 TCTAATATAGGTTTGGGCTTAAGAGAGTCCGAGCTTGTCCGAGAGGCGCTTAACAC 7680
Dh 7961 TCTAATATAGGTTTGGGCTTAAGAGAGTCCGAGCTTGTCCGAGAGGCGCTTAACAC 8020
Qy 7681 ATCAAGTCCGTGTGAAGAGCTCTCTGAGAGCTCAGAAAACCAATTCCCAACCAATT 7740
Dh 8021 ATCAAGTCCGTGTGAAGAGCTCTCTGAGAGCTCAGAAAACCAATTCCCAACCAATT 8080
Qy 7741 ATGGCCAAAATGAGGTCTTCTGAGTGAAGCCCAACCAAGGGGGCAAGAAAGAGCTGCG 7800
Dh 8081 ATGGCCAAAATGAGGTCTTCTGAGTGAAGCCCAACCAAGGGGGCAAGAAAGAGCTGCG 8140
Qy 7801 CTATCGTTTACCTGACCTCGGCGTCAAGGCTCTGAGAGATGCGCTTTATGACATT 7860
Dh 8141 CTATCGTTTACCTGACCTCGGCGTCAAGGCTCTGAGAGATGCGCTTTATGACATT 8200
Qy 7861 ACACAAAAACCTTCTCAGGCGGTGATGAGGGGCTTTATGATTCAGTATCCCGGCT 7920
Dh 8201 ACACAAAAACCTTCTCAGGCGGTGATGAGGGGCTTTATGATTCAGTATCCCGGCT 8260
Qy 7921 CAGCGGTGAGTTCCTCTGAAAAGCATGGGCGGAAAAGAGGACCTTAAGGGTTTTTCG 7980
Dh 8261 CAGCGGTGAGTTCCTCTGAAAAGCATGGGCGGAAAAGAGGACCTTAAGGGTTTTTCG 8320
Qy 7981 TATGATACCCGATGCTTGACTCAACCGTCACTGAGAGAGACATCAGAGCTGAGAGTCC 8040
Dh 8321 TATGATACCCGATGCTTGACTCAACCGTCACTGAGAGAGACATCAGAGCTGAGAGTCC 8380
Qy 8041 ATATATCGGCTCTGCTCTTGGCCGAGAGAGCCCACTGCACTACCTGCTTAAGTCTGAG 8100
Dh 8381 ATATATCGGCTCTGCTCTTGGCCGAGAGAGCCCACTGCACTACCTGCTTAAGTCTGAG 8440
Qy 8101 AGACTTTAAGTGGAGGGGCTTAAGTTCAACAGCAGAGGGCCAAACCTGCGGGTACAGGCT 8160
Dh 8441 AGACTTTAAGTGGAGGGGCTTAAGTTCAACAGCAGAGGGCCAAACCTGCGGGTACAGGCT 8500
Qy 8161 TGCAGGCGCAGCGGGGTGTCAACACTAGCATGGGGAACAACATCACTAGCTAGTGA 8220
Dh 8501 TGCAGGCGCAGCGGGGTGTCAACACTAGCATGGGGAACAACATCACTAGCTAGTGA 8560
Qy 8221 GCTTACGGCTTGTAAAGTGTGAGAGATATGCGCCCAATGCTGTATGCGGCGAT 8280
Dh 8561 GCTTACGGCTTGTAAAGTGTGAGAGATATGCGCCCAATGCTGTATGCGGCGAT 8620
Qy 8281 GACTGGTGTCTATCAGAAAAGCAGGGGACGAGAGAGACGAGCGGAACTGAGAGCC 8340
Dh 8621 GACTGGTGTCTATCAGAAAAGCAGGGGACGAGAGAGACGAGCGGAACTGAGAGCC 8680
Qy 8341 TTCACGAGAGCTATGACAGAGTATCTGCGCTCTGTGATACCCCGCAGACCGAGTAT 8400
Dh 8681 TTCACGAGAGCTATGACAGAGTATCTGCGCTCTGTGATACCCCGCAGACCGAGTAT 8740
Qy 8401 GATCTGAGCTGATTAATCTTGTCTCTCAAAATGTGTGTGTGGGCTGTGGGCTCCAAAGGC 8460
Dh 8741 GATCTGAGCTGATTAATCTTGTCTCTCAAAATGTGTGTGTGGGCTGTGGGCTCCAAAGGC 8800

QY 8461 CGCCGAGATACCTGACGAGACCTTACCTCAATCGCCGCGGCTGCTGGGAA 8520
PI CGCCGAGATACCTGACGAGACCTTACCTCAATCGCCGCGGCTGCTGGGAA 8860
DB 8801 CGCCGAGATACCTGACGAGACCTTACCTCAATCGCCGCGGCTGCTGGGAA 8860
QY 8521 ACAGTTAGACACTCCCTGTTGATTCATGCTGGGAAACATCATCCAGTACGCCGAGC 8580
XX ACAGTTAGACACTCCCTGTTGATTCATGCTGGGAAACATCATCCAGTACGCCGAGC 8580
DB 8861 ACAGTTAGACACTCCCTGTTGATTCATGCTGGGAAACATCATCCAGTACGCCGAGC 8920
QY 8581 ATATGGGCTCGCATGCTGCTGATGACACACTTTCTTCTCATTTCTCATGGCTCAAGACAG 8640
DB 8921 ATATGGGCTCGCATGCTGCTGATGACACACTTTCTTCTCATTTCTCATGGCTCAAGACAG 8980
QY 8641 CTGACACGAACTCTCAACTTTGAGATGTACGAGCGGTGTACTCCGTGAGTCCCTTGAC 8700
DB 8981 CTGACACGAACTCTCAACTTTGAGATGTACGAGCGGTGTACTCCGTGAGTCCCTTGAC 9040
QY 8701 CTCCAGCTATATTTGAAAGGTTACATGCGGCTTACCGCTTTTCTTGACACATACACT 8760
DB 9041 CTCCAGCTATATTTGAAAGGTTACATGCGGCTTACCGCTTTTCTTGACACATACACT 9100
QY 8761 CCCCACGAACTGACACGCGGTGCTTACGACCTCAGAAAATTGGGCGCCACCCCTGAC 8820
DB 9101 CCCCACGAACTGACACGCGGTGCTTACGACCTCAGAAAATTGGGCGCCACCCCTGAC 9160
QY 8821 GCGTGAAGAGCCGCGGACCTGTCAGTACGAGCGGCTCCCTCATCTCCGTGGGCGGAGAGC 8880
DB 9161 GCGTGAAGAGCCGCGGACCTGTCAGTACGAGCGGCTCCCTCATCTCCGTGGGCGGAGAGC 9220
QY 8881 GCGGTTTGGGCTGATATCTTTCAATTTGGGCGGTGAAGACCAAGCTCAAACTACTCCA 8940
DB 9221 GCGGTTTGGGCTGATATCTTTCAATTTGGGCGGTGAAGACCAAGCTCAAACTACTCCA 9280
QY 8941 TTGCGGAAAGCGCGCTCTGATTTTATTCAGTGTTCACGCTGCGCGCGGCGGCGGCG 9000
DB 9281 TTGCGGAAAGCGCGCTCTGATTTTATTCAGTGTTCACGCTGCGCGCGGCGGCGGCG 9340
QY 9001 GACATTTATCAACAGCGTGTGCGTGTGCGGACCCCGCTTATTTGCTTTGCTTACTCTTA 9060
DB 9341 GACATTTATCAACAGCGTGTGCGTGTGCGGACCCCGCTTATTTGCTTTGCTTACTCTTA 9400
QY 9061 CTTTGTGAGGGGTGAGGCTTTCTTCTTACTCCCGCTGCG 9099
DB 9401 CTTTGTGAGGGGTGAGGCTTTCTTCTTACTCCCGCTGCG 9439

RESULT 2
AAC86644
ID AAC86644 standard; DNA; 9711 BP.
XX AAC86644;
AC AAC86644;
XX 02-APR-2001 (first entry)
DT 02-APR-2001 (first entry)
DE Nucleotide sequence of infectious Hepatitis C virus strain HC-J6CH.
XX HC/J6CH; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine; ss.
KW Hepatitis C virus.
XX Hepatitis C virus.
FH Key Location/Qualifiers
FT 341..9442
FT CDS /product= "Hepatitis C virus protein"
XX MO200075338-A2.
XX 14-DEC-2000.
XX 02-JUN-2000; 2000MO-US015446.
XX 04-JUN-1999; 99US-0137693P.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Yanagi M, Bukh J, Emerson SU, Purcell RH;
PI WPI: 2001-061728/07.
DR P-PSDB; AAB30729.
XX DR
XX PT Nucleic acid molecule encoding human hepatitis C virus of genotype 2a for
PT developing vaccines, for diagnosis of hepatitis C virus and in screening
PT assays for identification of antiviral agents.
XX
PS Claim 3; Page 84-87; 167pp; English.
XX
XX The present sequence represents a nucleotide sequence of infectious
CC Hepatitis C virus (HCV) strain HC-J6CH genotype 2a. The HCV
CC polynucleotide sequence is capable of expressing the virus when
CC transfected into cells. The HCV protein is useful for assaying candidate
CC antiviral agents for activity against HCV. Antibodies specific for HCV
CC polypeptide are useful in prevention and treatment of diseases caused by
CC HCV in animals, in particular humans. The HCV polypeptides serve as
CC immunogens in the development of vaccines for preventing HCV in mammals
CC or as antigens in diagnostic assays for detecting the presence of HCV in
CC biological samples. The HCV polynucleotide is also useful for identifying
CC cell lines capable of supporting the replication of HCV in vitro and to
CC produce attenuated viral strains via passage in vitro or in vivo
XX
SQ Sequence 9711 BP; 1987 A; 2853 C; 2668 G; 2203 T; 0 U; 0 Other;
Query Match 100.0%; Score 9099; DB 5; Length 9711;
Best local similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;
Matches 9099; Conservative 0; Indels 0; Gaps 0;
QY 1 ATGAGCAAAATCTTAACTCAAGAAACCAAAAGAACCAACCGTCCGCCA 60
DB 341 ATGAGCAAAATCTTAACTCAAGAAACCAAAAGAACCAACCGTCCGCCA 400
QY 61 GACGTTAAGTTTCCGGGCGCGGCGGATCGTTGGGAGTATCTTTGGCCGGCAG 120
DB 401 GACGTTAAGTTTCCGGGCGCGGCGGATCGTTGGGAGTATCTTTGGCCGGCAG 460
QY 121 GGGCCAGGTTGGGTTGCGCGGACCAAGAAAGACTTGGAGCGGCTCCAGCACTGGA 180
DB 461 GGGCCAGGTTGGGTTGCGCGGACCAAGAAAGACTTGGAGCGGCTCCAGCACTGGA 520
QY 181 AGGCGCCAGCCATCCCTTAAAGATCGGCTCCACTGAGCAATCTTGGGAAAAACAGGA 240
DB 521 AGGCGCCAGCCATCCCTTAAAGATCGGCTCCACTGAGCAATCTTGGGAAAAACAGGA 580
QY 241 TACCCCTGGGCGCTTAAAGGAAATGAGGACTCGGTGGGCGAGATGGCTCTGCTCC 300
DB 581 TACCCCTGGGCGCTTAAAGGAAATGAGGACTCGGTGGGCGAGATGGCTCTGCTCC 640
QY 301 CGAGGTTCCCGGCTCTTGGGCGCCCAATGACCCCGGCAATAGTCCGCAACGTGGGT 360
DB 641 CGAGGTTCCCGGCTCTTGGGCGCCCAATGACCCCGGCAATAGTCCGCAACGTGGGT 700
QY 361 AAGGTATGATATCCCTTAAAGTGTGCGGCTTGGGACCTATAGGAGTATCTCTGTCTG 420
DB 701 AAGGTATGATATCCCTTAAAGTGTGCGGCTTGGGACCTATAGGAGTATCTCTGTCTG 760
QY 421 GGGCGCCCGCTCGGCGGCGTGGCAGAGCTTTGCGGCAATGGGTGAGATCTCTGGAGAG 480
DB 761 GGGCGCCCGCTCGGCGGCGTGGCAGAGCTTTGCGGCAATGGGTGAGATCTCTGGAGAG 820
QY 481 GGGGTTAATTTTGAACAGGAACTTACCGGTTGCTCTTTCTTCTTCTTCTGAGCC 540
DB 821 GGGGTTAATTTTGAACAGGAACTTACCGGTTGCTCTTTCTTCTTCTTCTGAGCC 880
QY 541 CTGCTGTCTGATCAACACCCCGATCTCGCTGCGGAAAGTAAAGAACTACGTACCGGC 600
DB 881 CTGCTGTCTGATCAACACCCCGATCTCGCTGCGGAAAGTAAAGAACTACGTACCGGC 940
QY 601 TACATGTTGATCAACAGCTGACCAATGACAGATTAAGTGGAGCTCCAGGCTGCTGTC 660

Db 941 TACATGGTGAATAACGACTGCAACCAATGACAGCATTAATCTGGAGCTCCAGGCTGCTGTC 1000
Qy 661 CTCCACGTCCTCCCGGGTGGCTTCCCGTGGAGAAAGTGGGGAATGATATTCAGTGTGGATA 720
Db 1001 CTCAACGTCCTCCCGGGTGGCTTCCCGTGGAGAAAGTGGGGAATGATATTCAGTGTGGATA 1060
Qy 721 CCGGCTCAACGAAATGAGCCGTGACAGCGGCGCCGCTCAACGAGGGCTTGGAGAG 780
Db 1061 CCGGCTCAACGAAATGAGCCGTGACAGCGGCGCCGCTCAACGAGGGCTTGGAGAG 1120
Qy 781 CACATGCAATGAGTGTGATGTCGCAACGCTCTGCTCTGCTCTCACTGAGGGGACCTC 840
Db 1121 CACATGCAATGAGTGTGATGTCGCAACGCTCTGCTCTGCTCTCACTGAGGGGACCTC 1180
Qy 841 TGGGGTGGGGTGAATGCTCCGACGCCAAATGTTCAATGTCCTCCGCGACCACTGGTTT 900
Db 1181 TGGGGTGGGGTGAATGCTCCGACGCCAAATGTTCAATGTCCTCCGCGACCACTGGTTT 1240
Qy 901 GTCGAAGACTGCAATGTCCTCATCTACCTGTGTACATCACTGGACACCGCATGGCATYGG 960
Db 1241 GTCGAAGACTGCAATGTCCTCATCTACCTGTGTACATCACTGGACACCGCATGGCATYGG 1300
Qy 961 GACATGATGATGAATCTGTCGCCCAACGCTACATGATCTTGGCGTACGCGATGCTGTC 1020
Db 1301 GACATGATGATGAATCTGTCGCCCAACGCTACATGATCTTGGCGTACGCGATGCTGTC 1360
Qy 1021 CCGGAGGTCAATTAAGACATCAATTAAGCGGGCTCATTTGGGGCTCATGTTCCGTTGGCC 1080
Db 1361 CCGGAGGTCAATTAAGACATCAATTAAGCGGGCTCATTTGGGGCTCATGTTCCGTTGGCC 1420
Qy 1081 TACTTCTATGACGGAGGCGTGGCGAAAGTGGTGTCACTCTTCTGTTGGCGCGGG 1140
Db 1421 TACTTCTATGACGGAGGCGTGGCGAAAGTGGTGTCACTCTTCTGTTGGCGCGGG 1480
Qy 1141 GTGAGCGCGGCAACCACTACTGTTGGGGGTTCTGCGCGGACACACCGGGCGCTCAC 1200
Db 1481 GTGAGCGCGGCAACCACTACTGTTGGGGGTTCTGCGCGGACACACCGGGCGCTCAC 1540
Qy 1201 AGCTTATTTGAACATGGGCCCCCAGGACAGAAATCCAGCTCGTTAAACCAATGGAGCTGG 1260
Db 1541 AGCTTATTTGAACATGGGCCCCCAGGACAGAAATCCAGCTCGTTAAACCAATGGAGCTGG 1600
Qy 1261 CACATCAACCGCAACCGGCTGAACTGCAATGACTCTTGGACACCGGCTTTATGCGCTC 1320
Db 1601 CACATCAACCGCAACCGGCTGAACTGCAATGACTCTTGGACACCGGCTTTATGCGCTC 1660
Qy 1321 CTGTTCTACACCCACAGCTTCAACTCGTCAGATGTCGCGAAGCATGTCGCGCTGCGC 1380
Db 1661 CTGTTCTACACCCACAGCTTCAACTCGTCAGATGTCGCGAAGCATGTCGCGCTGCGC 1720
Qy 1381 AGTATGAGGCTTCCGGGTGGGATGGGGCGCTTGGCAATTAAGAGTAAATGTCAACCAAT 1440
Db 1721 AGTATGAGGCTTCCGGGTGGGATGGGGCGCTTGGCAATTAAGAGTAAATGTCAACCAAT 1780
Qy 1441 CCAGAGGATTAAGAACCTTATGCTGGCACTAACCAACAGAGAGTGGCGTGTCTCC 1500
Db 1781 CCAGAGGATTAAGAACCTTATGCTGGCACTAACCAACAGAGAGTGGCGTGTCTCC 1840
Qy 1501 GCGAAGACTGTGTGTGGCCCAAGTGTACTGTTTCAACCCCAAGCCCAAGTGTGTGGAGCG 1560
Db 1841 GCGAAGACTGTGTGTGGCCCAAGTGTACTGTTTCAACCCCAAGCCCAAGTGTGTGGAGCG 1900
Qy 1561 ACCGACAGGCTTGAAGGCCCACTTAACGTTGGGGGAGAAATGAGAAAGTGTCTTCTA 1620
Db 1901 ACCGACAGGCTTGAAGGCCCACTTAACGTTGGGGGAGAAATGAGAAAGTGTCTTCTA 1960
Qy 1621 TTGAACAGCACTGACCAACGCTGGGGTCAATGTTCCGCTGACAGTGTGAATCTTCT 1680
Db 1961 TTGAACAGCACTGACCAACGCTGGGGTCAATGTTCCGCTGACAGTGTGAATCTTCT 2020
Qy 1681 GGTCAACCAAGACTTGGCGGACCAACCTGCGGTAATGAAGTGAATTTCAACGCGAC 1740
Db 2021 GGTCAACCAAGACTTGGCGGACCAACCTGCGGTAATGAAGTGAATTTCAACGCGAC 2080

Qy 1741 ACCGACCTGTTGGCCCAACGACTGTTTAAAGAGCATCTGATACCACTTACCTCAAA 1800
Db 2081 ACCGACCTGTTGGCCCAACGACTGTTTAAAGAGCATCTGATACCACTTACCTCAAA 2140
Qy 1801 TGGGCTCTGGGCTTGGCTTGGCTCAACGCAAGGTGCTGATCACTAACCCCTACAGCTTGG 1860
Db 2141 TGGGCTCTGGGCTTGGCTTGGCTCAACGCAAGGTGCTGATCACTAACCCCTACAGCTTGG 2200
Qy 1861 CATTAACCCCTGCAACGTTAACTATACATCTTCAAAATAGAGATGATATGAGAGGGGTT 1920
Db 2201 CATTAACCCCTGCAACGTTAACTATACATCTTCAAAATAGAGATGATATGAGAGGGGTT 2260
Qy 1921 GAGCAGAGCTCAACGCTGATGCAATTTCACTGAGGGAGTGGTTCATCTTGGAGAGAC 1980
Db 2261 GAGCAGAGCTCAACGCTGATGCAATTTCACTGAGGGAGTGGTTCATCTTGGAGAGAC 2320
Qy 1981 AGAGACAGAAAGTCAACTGTCCTCTTGTGTGCACTTCCACACGGAATGGGCCATTTTACT 2040
Db 2321 AGAGACAGAAAGTCAACTGTCCTCTTGTGTGCACTTCCACACGGAATGGGCCATTTTACT 2380
Qy 2041 TGCCTTATCTGGGACCTGCGGCTTGTGCACTGCTTCTTCCACTTCCACCAAAACATC 2100
Db 2381 TGCCTTATCTGGGACCTGCGGCTTGTGCACTGCTTCTTCCACTTCCACCAAAACATC 2440
Qy 2101 GTGGAAGTCAATTAATGATGAGCTTATCACTGACCTGACCAAAATACATGTCGATGG 2160
Db 2441 GTGGAAGTCAATTAATGATGAGCTTATCACTGACCTGACCAAAATACATGTCGATGG 2500
Qy 2161 GAGTGGTAAATTAATCTTATCTGCTTGTAGCGGACCGCAGGGTTTGGCTGTTATGG 2220
Db 2501 GAGTGGTAAATTAATCTTATCTGCTTGTAGCGGACCGCAGGGTTTGGCTGTTATGG 2560
Qy 2221 ATGCTATCTTGTGGGCCAGGCCGAAAGCAGCATGAGAAAGCTGCTTGGCAAGCTT 2280
Db 2561 ATGCTATCTTGTGGGCCAGGCCGAAAGCAGCATGAGAAAGCTGCTTGGCAAGCTT 2620
Qy 2281 GCGAGGCGACGTAGCTGCAATGAGCTTCTATATTTTGTCACTTTTGTGAGCTGCTGG 2340
Db 2621 GCGAGGCGACGTAGCTGCAATGAGCTTCTATATTTTGTCACTTTTGTGAGCTGCTGG 2680
Qy 2341 TACATCAAGGCTGGGATGATCCCTTAGCTAATCTTCCCTCACTGGCTGTGTCTTT 2400
Db 2681 TACATCAAGGCTGGGATGATCCCTTAGCTAATCTTCCCTCACTGGCTGTGTCTTT 2740
Qy 2401 AGCTTACTGCTCTAGCATTTGCCCAACAGGCTTATGCTTATGACGATCTGTGCAATGGC 2460
Db 2741 AGCTTACTGCTCTAGCATTTGCCCAACAGGCTTATGCTTATGACGATCTGTGCAATGGC 2800
Qy 2461 CAGATGAGAGGCTGCTGCTGATGATGATCACTCTTTACTCTACACCCCGGATTAAG 2520
Db 2801 CAGATGAGAGGCTGCTGCTGATGATGATCACTCTTTACTCTACACCCCGGATTAAG 2860
Qy 2521 ACCCTTCTCAACCGGTTTTGTGTGGTGTGTGATCTTCTGACCTTGGGGAAAGCTAAG 2580
Db 2861 ACCCTTCTCAACCGGTTTTGTGTGGTGTGTGATCTTCTGACCTTGGGGAAAGCTAAG 2920
Qy 2581 GTCCAGAGATGGGACCACTATGACAGTGGCGGGTGGCCGTGATGAGCATATATGGGCC 2640
Db 2921 GTCCAGAGATGGGACCACTATGACAGTGGCGGGTGGCCGTGATGAGCATATATGGGCC 2980
Qy 2641 GTGCGCAATTTTCAACAGGCTGTGTGTGATGATTAACAAAGTGGCTTGGCGGCTT 2700
Db 2981 GTGCGCAATTTTCAACAGGCTGTGTGTGATGATTAACAAAGTGGCTTGGCGGCTT 3040
Qy 2701 GGGCTGCTTAACTCTTAAAGGTGCTTGTGACCGCGTGTGCTTGTGCAAGGCTCAC 2760
Db 3041 GGGCTGCTTAACTCTTAAAGGTGCTTGTGACCGCGTGTGCTTGTGCAAGGCTCAC 3100
Qy 2761 GCTCTACTGAGATATGACATATGACAGGCAATCTGCGGGGGGAGAGTACATGTCAGATG 2820
Db 3101 GCTCTACTGAGATATGACATATGACAGGCAATCTGCGGGGGGAGAGTACATGTCAGATG 3160

QY 2821 GCAGTACTAGCCCTTGGAGGTGAGCTGACATTACATCTATGACCACTCACCCCTATG 2880
Db 3161 GCGCTACTAGCCCTTGGAGGTGAGCTGACATTACATCTATGACCACTCACCCCTATG 3220
QY 2881 TCGGATTTGGGCTGCTAGTGGCTCTGGGGACCTGGCCGCTCGCCCTTGAAGCCTATATCTTC 2940
Db 3221 TCGGATTTGGGCTGCTAGTGGCTCTGGGGACCTGGCCGCTCGCCCTTGAAGCCTATATCTTC 3280
QY 2941 AGTCCGATGAGAGAAAGAAATCATTTGCTGGGGAGCGGAGACGTGCTTGGGGACATT 3000
Db 3281 AGTCCGATGAGAGAAAGAAATCATTTGCTGGGGAGCGGAGACGTGCTTGGGGACATT 3340
QY 3001 TTACACGAGCTTCCCGTGTCCGCCGACCTTGTCTGGAGAGTCTCTCTTGGCCACGTGAT 3060
Db 3341 TTACACGAGCTTCCCGTGTCCGCCGACCTTGTCTGGAGAGTCTCTCTTGGCCACGTGAT 3400
QY 3061 GGGCTATACCTCCAGAGGGGTGGAGTCTTGGCCGCCACATCTGCTTAAAGCCACAGACACA 3120
Db 3401 GGGCTATACCTCCAGAGGGGTGGAGTCTTGGCCGCCACATCTGCTTAAAGCCACAGACACA 3460
QY 3121 CGTGGCCCTTTTGGGACACCATAGTGGTGAAGCATGACGGGGCGCGACAAGAACAGAGCT 3180
Db 3461 CGTGGCCCTTTTGGGACACCATAGTGGTGAAGCATGACGGGGCGCGACAAGAACAGAGCT 3520
QY 3181 GGGGAAATTCAGAGTCTGTCCACAGTCACTAGTCTTCTCGGAAACATCCATCTCGGGG 3240
Db 3521 GGGGAAATTCAGAGTCTGTCCACAGTCACTAGTCTTCTCGGAAACATCCATCTCGGGG 3580
QY 3241 GTTTTGTGGAGCTGTCTACCATGGAGCTGGCAACAAGACTGTGGCGGGCTCACGGGGTCCG 3300
Db 3581 GTTTTGTGGAGCTGTCTACCATGGAGCTGGCAACAAGACTGTGGCGGGCTCACGGGGTCCG 3640
QY 3301 GTCAACGAGATGTACTCCAGTGTCTGAGGGGAGCTTAAAGGTGGCCAGCCCTCTGGG 3360
Db 3641 GTCAACGAGATGTACTCCAGTGTCTGAGGGGAGCTTAAAGGTGGCCAGCCCTCTGGG 3700
QY 3361 ACTAAATCTTTGAGAGCCGTGCACTGTGTGAGCGGTGCACTGTACCTGTGTACCGCGAAC 3420
Db 3701 ACTAAATCTTTGAGAGCCGTGCACTGTGTGAGCGGTGCACTGTGTACCTGTGTACCGCGAAC 3760
QY 3421 GCTGATGTCACTCCGGCTCGAAGAGCGGGGAGCAAAAGGGGAGGCTACTCTCCCGAGA 3480
Db 3761 GCTGATGTCACTCCGGCTCGAAGAGCGGGGAGCAAAAGGGGAGGCTACTCTCCCGAGA 3820
QY 3481 CCTCTTTTCCACCTTGAAGGGGTCTCAGAGAGCCCGGTGCTATGCCCGAGGGGCAAGCT 3540
Db 3821 CCTCTTTTCCACCTTGAAGGGGTCTCAGAGAGCCCGGTGCTATGCCCGAGGGGCAAGCT 3880
QY 3541 GTCGGAGTCTTCCGGGCAAGCTGTGTCTCTCGGGGCGTGGCTAAGTCAATAGATTTTATC 3600
Db 3881 GTCGGAGTCTTCCGGGCAAGCTGTGTCTCTCGGGGCGTGGCTAAGTCAATAGATTTTATC 3940
QY 3601 CCGGTGAGAGCACTGACATGCTGACGGGGTCCCGACCTTAAAGCAACAGGACCA 3660
Db 3941 CCGGTGAGAGCACTGACATGCTGACGGGGTCCCGACCTTAAAGCAACAGGACCA 4000
QY 3661 CCTGCTGTGCCCCAGACCTATCAGGTCCGGTACTTTCATGCTGACGCCCCGACTGGCAGTGAAG 3720
Db 4001 CCTGCTGTGCCCCAGACCTATCAGGTCCGGTACTTTCATGCTGACGCCCCGACTGGCAGTGAAG 4060
QY 3721 AGCACCAAGTCTCTGTGCAATGCTGTCAAGGGGTAAAGTGTAGTGTCTTATCTCC 3780
Db 4061 AGCACCAAGTCTCTGTGCAATGCTGTCAAGGGGTAAAGTGTAGTGTCTTATCTCC 4120
QY 3781 TCAGTGGCTGCCACCTGGGGTTTGGGGGTAACTTGTCTAAGGACATGGGATCAATCC 3840
Db 4121 TCAGTGGCTGCCACCTGGGGTTTGGGGGTAACTTGTCTAAGGACATGGGATCAATCC 4180
QY 3841 AACATTAGAGACTGAGTCAAGACTGTGACACCGGGGCGCCATCAGCTACCAATAT 3900
Db 4181 AACATTAGAGACTGAGTCAAGACTGTGACACCGGGGCGCCATCAGCTACCAATAT 4240
QY 3901 GGGAAATTCCTGCGGATGGGGGCTGTGGCGGGCGGCTTACGACATCATATGTGAT 3960

Db 4241 GGGAAATTCCTGCGGATGGGGGCTGTGGCGGGCGGCTTACGACATCATATGTGAT 4300
QY 3961 GAATGCCATGCGGTGAGCTCTACCAACATCTTGGGATCGAAAGTCTTATGAAGA 4020
Db 4301 GAATGCCATGCGGTGAGCTCTACCAACATCTTGGGATCGAAAGTCTTATGAAGA 4360
QY 4021 GAGACAGTGGGGTGAAGCTAACTGTGTGCTAGCTAACGCTACGCCCCCTGGGTCAGTGACA 4080
Db 4361 GAGACAGTGGGGTGAAGCTAACTGTGTGCTAGCTAACGCTACGCCCCCTGGGTCAGTGACA 4420
QY 4081 ACCCCCAACCCCACTAAGAGAGTGGCCCTTGGGCAAGAGGGGAGATCCCTCTAT 4140
Db 4421 ACCCCCAACCCCACTAAGAGAGTGGCCCTTGGGCAAGAGGGGAGATCCCTCTAT 4480
QY 4141 GGGAGGGGATTCCTCTGTCTTACATCAAGGAGAGAAAGATCTGATCTTGCATTTCA 4200
Db 4481 GGGAGGGGATTCCTCTGTCTTACATCAAGGAGAGAAAGATCTGATCTTGCATTTCA 4540
QY 4201 AAGAAAAAGTGTGACGAGCTCGCGGGCCCTTTCGGGGTATGGGCTTGAATCAGTGACA 4260
Db 4541 AAGAAAAAGTGTGACGAGCTCGCGGGCCCTTTCGGGGTATGGGCTTGAATCAGTGACA 4600
QY 4261 TACTACAGAGGTGAGAGCTTCCCTAATACCAACTCAGGGAGAGCTAGTGTCTGCGCC 4320
Db 4601 TACTACAGAGGTGAGAGCTTCCCTAATACCAACTCAGGGAGAGCTAGTGTCTGCGCC 4660
QY 4321 ACCGAGCCCTCATGACAGGAGTAACTGAGGGGACTTGTGACTCGGTATGCACTGACACGTA 4380
Db 4661 ACCGAGCCCTCATGACAGGAGTAACTGAGGGGACTTGTGACTCGGTATGCACTGACACGTA 4720
QY 4381 GCGGTCACTCAAGTTGTAGACTTCAAGTTTAAAGCCCACTTCACTAACCACAGATT 4440
Db 4721 GCGGTCACTCAAGTTGTAGACTTCAAGTTTAAAGCCCACTTCACTAACCACAGATT 4780
QY 4441 GTCCCTCAAGAGCTGTCTCAGTACAGCCCGGGGGTGGACGGGTGAGGGGAAAGACTG 4500
Db 4781 GTCCCTCAAGAGCTGTCTCAGTACAGCCCGGGGGTGGACGGGTGAGGGGAAAGACTG 4840
QY 4501 GGCATTTATAGGATGTTTCACTGGTGAAGCGAGCTCAGAGATGTTTGAAGTGTAGT 4560
Db 4841 GGCATTTATAGGATGTTTCACTGGTGAAGCGAGCTCAGAGATGTTTGAAGTGTAGT 4900
QY 4561 CTCTGTGAGTGTACGACGAGGGGCGGATGGTATGAGTCAACATCGAGACACC 4620
Db 4901 CTCTGTGAGTGTACGACGAGGGGCGGATGGTATGAGTCAACATCGAGACACC 4960
QY 4621 GTGAGGCTCAGGGCGGATTTCAACAGCCCGGTTGCTGTGTGCGCAAGACATCTTAG 4680
Db 4961 GTGAGGCTCAGGGCGGATTTCAACAGCCCGGTTGCTGTGTGCGCAAGACATCTTAG 5020
QY 4681 TTTTGGGAGGCAAGTTTCAACCGGCTCAGACATAGATGCCACTTCTTCCAAACA 4740
Db 5021 TTTTGGGAGGCAAGTTTCAACCGGCTCAGACATAGATGCCACTTCTTCCAAACA 5080
QY 4741 AAGCAATCGGGGAAAAATTTGCACTTAACAGCTTACAGGCTACAGGTGCGCTAAG 4800
Db 5081 AAGCAATCGGGGAAAAATTTGCACTTAACAGGCTTACAGGCTACAGGTGCGCTAAG 5140
QY 4801 GCGAAAGCCCCCCCCCGTCTGGGAGTCAATGTGAAGTGTGACTGAACTCAAGGCC 4860
Db 5141 GCGAAAGCCCCCCCCCGTCTGGGAGTCAATGTGAAGTGTGACTGAACTCAAGGCC 5200
QY 4861 ACACTGTGGGCCCCACACTCTCTGTACCGGTTGGGCTGTGTACAAAGAGTCAAC 4920
Db 5201 ACACTGTGGGCCCCACACTCTCTGTACCGGTTGGGCTGTGTACAAAGAGTCAAC 5260
QY 4921 CTCACATCCGCTGACGAATATCATGCGCACTGATGAGCCGACCTTGAAGTCAATG 4980
Db 5261 CTCACATCCGCTGACGAATATCATGCGCACTGATGAGCCGACCTTGAAGTCAATG 5320
QY 4981 ACCAGACATGGGTCTTGGCAGGGGAGTGTGGCGGCGGTGCGGATTTGGCTGGGG 5040

Dh 5321 ACCAGCACAATGGGCTTTGGCAGGGGAGTCTTGCGGCGCTGCGCGGTATGCTTGCG 5380
Qy 5041 ACCGGGTGTTTGGCATTCATCGGCGCTTGACATTTAACAGAGAGCGCGTGTGGCCG 5100
Dh 5381 ACCGGGTGTTTGGCATTCATCGGCGCTTGACATTTAACAGAGAGCGCGTGTGGCCG 5440
Qy 5101 GACAAAGAGTCTCTATGAGGCTTTGATGATGAGAGATGTGCTCTAGGCGGCT 5160
Dh 5441 GACAAAGAGTCTCTATGAGGCTTTGATGATGAGAGATGTGCTCTAGGCGGCT 5500
Qy 5161 CTCATTGAAGAGGGGAGCGGAGTAGCCGAGATCTGAAGTCCAAGATCCAAGGCTTATTG 5220
Dh 5501 CTCATTGAAGAGGGGAGCGGAGTAGCCGAGATCTGAAGTCCAAGATCCAAGGCTTATTG 5560
Qy 5221 GAGCAGCTTCCAAACAGGTCAGACATTCACCCATCTGTGAGGTTTATGCCCCAG 5280
Dh 5561 CAGCAAGCTTCCAAACAGGTCAGACATTCACCCATCTGTGAGGTTTATGCCCCAG 5620
Qy 5281 GTAGAACATTTCTGGGCAACACATGTGGAATCTTATGCGGACATTCATCTGCA 5340
Dh 5621 GTAGAACATTTCTGGGCAACACATGTGGAATCTTATGCGGACATTCATCTGCA 5680
Qy 5341 GAGCTATCAACACTGCGCAGGGAACCTGAGTAGCTTCATGATGCGGTTCACTGCGCC 5400
Dh 5681 GAGCTATCAACACTGCGCAGGGAACCTGAGTAGCTTCATGATGCGGTTCACTGCGCC 5740
Qy 5401 CTCACAGTCCGCTGTCAACACACATTCCTTCTCAACATTTTGGGGGCTGCTGA 5460
Dh 5741 CTCACAGTCCGCTGTCAACACACATTCCTTCTCAACATTTTGGGGGCTGCTGA 5800
Qy 5461 GCATCCCAATTTGACACACCCGCGGGGACACTGCGCTTCTGTTCAGTGGCTAGTGGGA 5520
Dh 5801 GCATCCCAATTTGACACACCCGCGGGGACACTGCGCTTCTGTTCAGTGGCTAGTGGGA 5860
Qy 5521 GCTGCGGTAGGAGATATAGGCTTAGTAACTGCTAGTGGACATCTCGGAGGCTATGAT 5580
Dh 5861 GCTGCGGTAGGAGATATAGGCTTAGTAACTGCTAGTGGACATCTCGGAGGCTATGAT 5920
Qy 5581 GCGGGCATTTGCGGGGCTCTGCTGCGCATTCAGATCATGCTGTGCGAGAACCCCTCCATG 5640
Dh 5921 GCGGGCATTTGCGGGGCTCTGCTGCGCATTCAGATCATGCTGTGCGAGAACCCCTCCATG 5980
Qy 5641 GAGGATGTGCTCAACTGCTGCTGCGCATTCGCTCCGCGGCTCTGTGATGAGGAGTC 5700
Dh 5981 GAGGATGTGCTCAACTGCTGCTGCGCATTCGCTCCGCGGCTCTGTGATGAGGAGTC 6040
Qy 5701 ATCTGCGCGGCAATTTCTGCGCGACACGTGGGACCGGGGAGAGCGCGCTCAATGATG 5760
Dh 6041 ATCTGCGCGGCAATTTCTGCGCGACACGTGGGACCGGGGAGAGCGCGCTCAATGATG 6100
Qy 5761 AATAGACTATTGCTTTTGTGCTTCAGAGGAATCATGCTGCGCCCACTACCTGACG 5820
Dh 6101 AATAGACTATTGCTTTTGTGCTTCAGAGGAATCATGCTGCGCCCACTACCTGACG 6160
Qy 5821 GAGTGGAGTGCCTGCGAGGCTGACCCCACTGCTGCTGCTTACCATACAGGCTG 5880
Dh 6161 GAGTGGAGTGCCTGCGAGGCTGACCCCACTGCTGCTGCTTACCATACAGGCTG 6220
Qy 5881 CTCAGAGAATCTCACAACCTGATTAAGAGAGCTGCCCCATCCATGCGGCGCTGCTG 5940
Dh 6221 CTCAGAGAATCTCACAACCTGATTAAGAGAGCTGCCCCATCCATGCGGCGCTGCTG 6280
Qy 5941 CTCGCGAGTGTGGAGCTGGGTTTGACACATCTTAACAGACTTTAAATAATGCTGAC 6000
Dh 6281 CTCGCGAGTGTGGAGCTGGGTTTGACACATCTTAACAGACTTTAAATAATGCTGAC 6340
Qy 6001 TCCAAATTAATCCCAAGATGCGGCTGCTCCCTTTGTCTCTGTCAAAAAGGGGTACAAG 6060
Dh 6341 TCCAAATTAATCCCAAGATGCGGCTGCTCCCTTTGTCTCTGTCAAAAAGGGGTACAAG 6400
Qy 6061 GGGGTGTGGGCGGCACTGGGATCATGACACACAGGCTGCTTGGCGGCGCAATATCTCT 6120
Dh 6401 GGGGTGTGGGCGGCACTGGGATCATGACACACAGGCTGCTTGGCGGCGCAATATCTCT 6460

Qy 6121 GGCATGTGCGCTTTGGGCTTCATGAGATTCACGGGGCTTAAGACTGTCATGAATATCTGG 6180
Dh 6461 GGCATGTGCGCTTTGGGCTTCATGAGATTCACGGGGCTTAAGACTGTCATGAATATCTGG 6520
Qy 6181 CAGGGACCTTTCTCTATCAATTGTTTACAGGAGGAGCAGTGCCTGCGAAAACCCGCGCA 6240
Dh 6521 CAGGGACCTTTCTCTATCAATTGTTTACAGGAGGAGCAGTGCCTGCGAAAACCCGCGCA 6580
Qy 6241 AACTTTAAGTGCCTATCTGAGAGGTGGGCTCAGAGTACCGCGAGGTGAGCGAC 6300
Dh 6581 AACTTTAAGTGCCTATCTGAGAGGTGGGCTCAGAGTACCGCGAGGTGAGCGAC 6640
Qy 6301 GGGTCATACCATACATCAATACAGGACTCACCACTGATTAATTGAAAGTCCCTGCAACTA 6360
Dh 6641 GGGTCATACCATACATCAATACAGGACTCACCACTGATTAATTGAAAGTCCCTGCAACTA 6700
Qy 6361 CCTCTCCGAGTCTTTCTGAGTGAAGAGAGTCAATCCATAGTGTGCCCCACA 6420
Dh 6701 CCTCTCCGAGTCTTTCTGAGTGAAGAGAGTCAATCCATAGTGTGCCCCACA 6760
Qy 6421 CCGAAGCGTTTTTCCGGGATGAGTCTGTTCTGCGTTGGCTTAATTATTTGCTG 6480
Dh 6761 CCGAAGCGTTTTTCCGGGATGAGTCTGTTCTGCGTTGGCTTAATTATTTGCTG 6820
Qy 6481 GGGTCCAGCTCTTGGCAGCCCTGAACCGGACACAGAGTATGATGTCATGCTAACA 6540
Dh 6821 GGGTCCAGCTCTTGGCAGCCCTGAACCGGACACAGAGTATGATGTCATGCTAACA 6880
Qy 6541 GATCATCTCATATCAACGCGGAGAGCTGACGCGGCGTTTAGCGCGGGGTACACCCCA 6600
Dh 6881 GATCATCTCATATCAACGCGGAGAGCTGACGCGGCGTTTAGCGCGGGGTACACCCCA 6940
Qy 6601 TCCGAGGCAAGCTCTCTGCGGAGAGCTATGCGGACCATGCTGCGGAGGCTACCTGACC 6660
Dh 6941 TCCGAGGCAAGCTCTCTGCGGAGAGCTATGCGGACCATGCTGCGGAGGCTACCTGACC 7000
Qy 6661 ACCCAGGCAAGGCTATGATGAGCATGCTGATGAGCTTAACCTGTTCAATGGGGGAGAT 6720
Dh 7001 ACCCAGGCAAGGCTATGATGAGCATGCTGATGAGCTTAACCTGTTCAATGGGGGAGAT 7060
Qy 6721 GTGACTCGATAGAGTCTGGGTCCAAAGTGTGCTTCTGAGCTCTCTGACCCCAATGCTC 6780
Dh 7061 GTGACTCGATAGAGTCTGGGTCCAAAGTGTGCTTCTGAGCTCTCTGACCCCAATGCTC 7120
Qy 6781 GAAAGAAAGAGGCACTTGAAGCTTGCATACATCAAGATACATGCTCCCAAGAAAGG 6840
Dh 7121 GAAAGAAAGAGGCACTTGAAGCTTGCATACATCAAGATACATGCTCCCAAGAAAGG 7180
Qy 6841 TTCCGACAGCTTTTACCGGCTGGGACAGGCTGATTAACAACCAACGCTTGTGGAATCG 6900
Dh 7181 TTCCGACAGCTTTTACCGGCTGGGACAGGCTGATTAACAACCAACGCTTGTGGAATCG 7240
Qy 6901 TGGAAAGGCAAGATTACCAACCGGCACTGTTGCGGCTGTGCTCTCCCTCTCTAGG 6960
Dh 7241 TGGAAAGGCAAGATTACCAACCGGCACTGTTGCGGCTGTGCTCTCCCTCTCTAGG 7300
Qy 6961 AAAACCCGACGCTCTCCCAAGAGAGCGCGGACAGTGGGCTTAAGTGAAGCTCATTA 7020
Dh 7301 AAAACCCGACGCTCTCCCAAGAGAGCGCGGACAGTGGGCTTAAGTGAAGCTCATTA 7360
Qy 7021 GAGAGTGCCTTCAACAGCTGAGCATTAAGTCTTTTGGCAGGCGCCGCCCAAGCGGAGAT 7080
Dh 7361 GAGAGTGCCTTCAACAGCTGAGCATTAAGTCTTTTGGCAGGCGCCGCCCAAGCGGAGAT 7420
Qy 7081 TCAAGGCTTTTCAACGAGGAGCGGAGCTGCGATTCGCGGAGTCAAGCGCTCTGTATGAG 7140
Dh 7421 TCAAGGCTTTTCAACGAGGAGCGGAGCTGCGATTCGCGGAGTCAAGCGCTCTGTATGAG 7480
Qy 7141 TTGGCCCTTTTCCGAGCAGGCTTCATCTCTCCATGCCCCCTCGAGGGGAGGCTTTGGA 7200
Dh 7481 TTGGCCCTTTTCCGAGCAGGCTTCATCTCTCTCCATGCCCCCTCGAGGGGAGGCTTTGGA 7540

OY	7201	GATCGAAGCTCTGAGACCTGAGACAGGTAGAGCCCAACCCGCCCCCAAGGGGGGGGTGAGCA	7250
Db	7541	GATCCAGACCTGAGACCTGAGACAGGTAGAGCCCAACCCGCCCCCAAGGGGGGGGTGAGCA	7600
OY	7261	GCTCCCGGCTCGAGACTGGGGGTCTTGCTTACTTGCTCCGAGGAGAGCACTCCGTGCTG	7320
Db	7601	GCTCCCGGCTCGAGACTCGGGGGTCTCGGTCTACTTGCTCCGAGGAGAGCACTCCGTGCTG	7660
OY	7321	TGCTGCTCCATGTCACTACTCTCGAGACGGGGGCTCTAATACTCCGTGAGTCCGAAAGAG	7380
Db	7661	TGCTGCTCCATGTCACTACTCTCGAGACGGGGGCTCTAATACTCCGTGAGTCCGAAAGAG	7720
OY	7381	GAGAAAGTTACCGAATTAACCCCTTGAGCAACTCCCTGTGCGATATCACAACAGGTGTAC	7440
Db	7721	GAGAAAGTTACCGAATTAACCCCTTGAGCAACTCCCTGTGCGATATCACAAGGTGTAC	7780
OY	7441	TGTACCAACAACAAAGAGGCGCTCACTAAGGGCTAAAAAGTAACTTTTGAATGAGATGCA	7500
Db	7781	TGTACCAACAACAAAGAGGCGCTCACTAAGGGCTAAAAAGTAACTTTTGAATGAGATGCA	7840
OY	7501	GTGCTCGACTCTCTACTAGACTAGCTCTTAAGGACATTAAAGCTAAGCGGCTCCAAAGTC	7560
Db	7841	GTGCTCGACTCTCTACTAGACTAGCTCTTAAGGACATTAAAGCTAAGCGGCTCCAAAGTC	7900
OY	7561	ACCGCAAGGCTCTTCACCATGAGAGGCGTTGCGAGTTAACCCCACTTCGCAAGA	7620
Db	7901	ACCGCAAGGCTCTTCACCATGAGAGGCGTTGCGAGTTAACCCCACTTCGCAAGA	7960
OY	7621	TCTAAATATGAGGTTTGGGGGCTTAAGGAGTCCGAGCTGTGCGGGAGGGCGGTTAACAC	7680
Db	7961	TCTAAATATGAGGTTTGGGGGCTTAAGGAGTCCGAGCTGTGCGGGAGGGCGGTTAACAC	8020
OY	7681	ATCAAGTCGGTGTGAGAGGAGCTCTCTGAGGACTCAGAAACCAATTCCCAACAATT	7740
Db	8021	ATCAAGTCGGTGTGAGAGGAGCTCTCTGAGGACTCAGAAACCAATTCCCAACAATT	8080
OY	7741	ATGGCCAAAAATGAGGTGTCTTGCTGTGACCCCAACCAAGGGGGGCAAGAAAGCACTGCG	7800
Db	8081	ATGGCCAAAAATGAGGTGTCTTGCTGTGACCCCAACCAAGGGGGGCAAGAAAGCACTGCG	8140
OY	7801	CTTATCGTTTACCCCGAAGCTCGGCGTCAAGGGCTGTGCGAAGATGAGCCCTTATATGACATT	7860
Db	8141	CTTATCGTTTACCCCGAAGCTCGGCGTCAAGGGCTGTGCGAAGATGAGCCCTTATATGACATT	8200
OY	7861	ACACAAAACTCTCTGAGGCGGTGATGAGGCGCTTTATGGAATTCAGTATTTCCCGCGCT	7920
Db	8201	ACACAAAACTCTCTGAGGCGGTGATGAGGCGCTTTATGGAATTCAGTATTTCCCGCGCT	8260
OY	7921	CAGCGGGTAGAGTTTCTCTGTGAAACATGAGGCGGAAAGAAAGGAAACCTATGAGGTTTTTCG	7980
Db	8261	CAGCGGGTAGAGTTTCTCTGTGAAACATGAGGCGGAAAGAAAGGAAACCTATGAGGTTTTTCG	8320
OY	7981	TATGATACCCGATGCTTTGACTCAACCGTCACTGAGAGAGCATCAGGACTGAGGAGTCC	8040
Db	8321	TATGATACCCGATGCTTTGACTCAACCGTCACTGAGAGAGCATCAGGACTGAGGAGTCC	8380
OY	8041	ATATATCGGGCGCTGCTCTCTGCGCGAGAGAGGCCCACTGTGATCACTCCCTTAACGAG	8100
Db	8381	ATATATCGGGCGCTGCTCTCTGCGCGAGAGAGGCCCACTGTGATCACTCCCTTAACGAG	8440
OY	8101	AGACTTATAGTGGGAGGCGCTATATGTTCAACAGCAAGGCGCAAACTTGCGGGTACAGCGGT	8160
Db	8441	AGACTTATAGTGGGAGGCGCTATATGTTCAACAGCAAGGCGCAAACTTGCGGGTACAGCGGT	8500
OY	8161	TGCGCGCGACGCGGGGTGCTCAACACTAGCATGAGGAAACCATATCACTACTAGTGA	8220
Db	8501	TGCGCGCGACGCGGGGTGCTCAACACTAGCATGAGGAAACCATATCACTACTAGTGA	8560
OY	8221	GCCTTAGCGGCTTTGTTAAAGCTSCAAGGATTAATTCGCGCCCAATCTGTGTTATGCGGCAT	8280
Db	8561	GCCTTAGCGGCTTTGTTAAAGCTSCAAGGATTAATTCGCGCCCAATCTGTGTTATGCGGCAT	8620
OY	8281	GACTTGGTTGATCTCAGAAAGCCAGGAGGACCGAGAGGAGCGAGCGGAACTGTGAGGC	8340

Db	8621	GACCTGGTTCTCATCTCCAGAAAGCCAGGGGAGCCGAGAGGAGCCAGAGCCGAACTGTAGAGCC	8680
Qy	8341	TTCA CGAGGCGTATAGACCAAGTATTTCTGCCCTCTGTGTGACCCCGCCAGACCGAGTAT	8400
Db	8681	TTCA CGAGGCGTATAGACCAAGTATTTCTGCCCTCTGTGTGACCCCGCCAGACCGAGTAT	8740
Qy	8401	GATCTGGAAGGTATACATCTTGTGCTCTCAAAATGTCTGTGGGCGCTGGGGCCCAAGAGC	8460
Db	8741	GATCTGGAAGGTATACATCTTGTGCTCTCAAAATGTCTGTGGGCGCTGGGGCCCAAGAGC	8800
Qy	8461	CGCGCGAGATACACCTGACCAAGACCTTACCACTCAATTCGCCCGAGCTGCTGTGGAA	8520
Db	8801	CGCGCGAGATACACCTGACCAAGACCTTACCACTCAATTCGCCCGAGCTGCTGTGGAA	8860
Qy	8521	ACAAGTTAGACACTCCCTGTCAATTCAATGGCTGTGGAAACATCATCAGTACGCCCGACC	8580
Db	8861	ACAAGTTAGACACTCCCTGTCAATTCAATGGCTGTGGAAACATCATCAGTACGCCCGACC	8920
Qy	8581	ATATGGGCTGTGCATGTGTCTGTATGACACACTTTCTTCTCACTTGTATGGCTCAAGACAGC	8640
Db	8921	ATATGGGCTGTGCATGTGTCTGTATGACACACTTTCTTCTCACTTGTATGGCTCAAGACAGC	8980
Qy	8641	CTGGAACAGAAACCTCAACTTTGAGATGTATAGAGGCGGTATACTCGTAGTCCCTTGAGAC	8700
Db	8981	CTGGAACAGAAACCTCAACTTTGAGATGTATAGAGGCGGTATACTCGTAGTCCCTTGAGAC	9040
Qy	8701	CTCCGACCTATTAATTGAAAGTTATCATGTGGCTTGACGCTTTTCTCTTGACACATACACT	8760
Db	9041	CTCCGACCTATTAATTGAAAGTTATCATGTGGCTTGACGCTTTTCTCTTGACACATACACT	9100
Qy	8761	CCCCAGAACTGACACGGGTGTGCTTCAAGCCTCAGAAAATTGGGGCGCCACCCCTTGAGA	8820
Db	9101	CCCCAGAACTGACACGGGTGTGCTTCAAGCCTCAGAAAATTGGGGCGCCACCCCTTGAGA	9160
Qy	8821	GCGTGGAAAGCCGGGGACGTGTGACGTAGGGCGTCCCTCAATCCCGGTGGGGGGAGAGCG	8880
Db	9161	GCGTGGAAAGCCGGGGACGTGTGACGTAGGGCGTCCCTCAATCCCGGTGGGGGGAGAGCG	9220
Qy	8881	GCCTTTGCGGTGATATCTCTTCAATTGGGCGGTGAAAGACCAAGCTCAAACTCATCTCA	8940
Db	9221	GCCTTTGCGGTGATATCTCTTCAATTGGGCGGTGAAAGACCAAGCTCAAACTCATCTCA	9280
Qy	8941	TTGCGCGAAGCGCGCTCTCGATTTTATCCAGCTGTTCACCGTGGCGCGCGGCGGGAGC	9000
Db	9281	TTGCGCGAAGCGCGCTCTCGATTTTATCCAGCTGTTCACCGTGGCGCGCGGCGGGAGC	9340
Qy	9001	GACATTTATCAACAGCGTGTGCGGTGCCCGACCCCGCTTATGTGCTTTGGCTTACTCTTA	9060
Db	9341	GACATTTATCAACAGCGTGTGCGGTGCCCGACCCCGCTTATGTGCTTTGGCTTACTCTTA	9400
Qy	9061	CTTTTGTAGGGGTAGGCTTTTCTCTATCCCGGCTGG	9099
Db	9401	CTTTTGTAGGGGTAGGCTTTTCTCTATCCCGGCTGG	9439

	RESULT 3
AAAF23486	AAAF23486
ID	AAF23486 standard; DNA; 9711 BP.
XX	
AC	AAF23486;
XX	
DT	21-MAR-2001 (first entry)
XX	
DE	Infectious Hepatitis C virus 2a genotype
XX	
KM	GBV-B; hepatitis C virus; HCV; vaccine; d
XX	
OS	Hepatitis C virus.
XX	
PN	MO200075337-A1.
XX	
PD	14-DEC-2000.

14-DEC-2000.

XX 02-JUN-2000; 2000WO-US015293.
PF
XX 04-JUN-1999; 99US-0137694P.
PR
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA
XX Buhk J, Yanagi M, Emerson SU, Purcell RH;
PI WPI: 2001-091214/10.
XX
DR
XX New infectious nucleic acids of the GB virus-B clone, useful for
PT indirectly studying the molecular properties of hepatitis C virus (HCV)
PT and in developing vaccines and therapeutics for HCV.
XX
PS Disclosure; Page 78-82; 96pp; English.
XX
CC The present invention relates to GB virus-B. The nucleic acid molecules
CC of the invention are useful for indirectly studying the molecular
CC properties of hepatitis C virus (HCV). The infectious nucleic acid
CC sequence of the GB virus-B clone and the HCV/GBV-B chimeras may be used
CC in the development of vaccines and therapeutics for HCV
XX
SQ Sequence 9711 BF; 1989 A; 2851 C; 2668 G; 2203 T; 0 U; 0 Other;
Query Match 100.0%; Score 9095.8; DB 4; Length 9711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9097; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCACAATCTTAACTCAAGAAAACCAAAAGAAACACCAACCGTCGCCACA 60
DB 341 ATGAGCACAATCTTAACTCAAGAAAACCAAAAGAAACACCAACCGTCGCCACA 400
QY 61 GACGTTAAGTTCCGGGCGCGCGCGCAATCTTGGCGGAGTACTTGTTCGCCGCAAG 120
DB 401 GACGTTAAGTTCCGGGCGCGCGCGCAATCTTGGCGGAGTACTTGTTCGCCGCAAG 460
QY 121 GGCCCAAGTTGGGTGTGCGCGCAAGAAAGACTTCGAGCGGTCGCCAGCACTGGA 180
DB 461 GGCCCAAGTTGGGTGTGCGCGCAAGAAAGACTTCGAGCGGTCGCCAGCACTGGA 520
QY 181 AGGCGCAAGCCATCTCTAAAGATCGGCGCTCACTGGCAAAATCTGGGGAACCAAGA 240
DB 521 AGGCGCAAGCCATCTCTAAAGATCGGCGCTCACTGGCAAAATCTGGGGAACCAAGA 580
QY 241 TACCCCTGGCCCTTATCGGGAATGAGGAGCTGGGCTGGGAGATGGCTCTGTCGCC 300
DB 581 TACCCCTGGCCCTTATCGGGAATGAGGAGCTGGGCTGGGAGATGGCTCTGTCGCC 640
QY 301 CGAGGTCGCCGTCCTTGGGGGCGCAATGACCCCGGCAATAGGTGGCGCAAGTGGGT 360
DB 641 CGAGGTCGCCGTCCTTGGGGGCGCAATGACCCCGGCAATAGGTGGCGCAAGTGGGT 700
QY 361 AAGGTCATCGATACCTTAACGTGCGGCTTGGCGGACTCACTGAGGTAATCTCTGTG 420
DB 701 AAGGTCATCGATACCTTAACGTGCGGCTTGGCGGACTCACTGAGGTAATCTCTGTG 760
QY 421 GGCGCCCGCTCGCGCGCGCTGCGAGAGCTTCGCGCATGGCGTGAAGTCTTGAAGAC 480
DB 761 GGCGCCCGCTCGCGCGCGCTGCGAGAGCTTCGCGCATGGCGTGAAGTCTTGAAGAC 820
QY 481 GGAGTTAATTTTGAACAGGAACTTACCCGAGTGTCTTTTCTATCTTCTGTCGAGC 540
DB 821 GGAGTTAATTTTGAACAGGAACTTACCCGAGTGTCTTTTCTATCTTCTGTCGAGC 880
QY 541 CTGCTGTCTGCAATCAACCCCGGCTCTCGGCTGCGAAGTGAAGCAATCAATGACG 600
DB 881 CTGCTGTCTGCAATCAACCCCGGCTCTCGGCTGCGAAGTGAAGCAATCAATGACG 940
QY 601 TACATGTGATCACTGACCAATGAGAGCAATTAATCTGAGAGCTTCAGGCTGCTGC 660
DB 941 TACATGTGATCACTGACCAATGAGAGCAATTAATCTGAGAGCTTCAGGCTGCTGC 1000

QY 661 CTCACGTCGCCGAGTCCTCCGTCGAGAAAGTGGGAATGATCTCACTGTCGAGTA 720
DB 1001 CTCACGTCGCCGAGTCCTCCGTCGAGAAAGTGGGAATGATCTCACTGTCGAGTA 1060
QY 721 CCGGTCACCGAATGTGGCCGTGACAGCGCGCGCGCTTACGACGAGGCTTGGAGC 780
DB 1061 CCGGTCACCGAATGTGGCCGTGACAGCGCGCGCGCTTACGACGAGGCTTGGAGC 1120
QY 781 CACATGCAATGATGATGATGTCGCGACGCTGCTGCGCTTCACTGATGAGGAGACCTC 840
DB 1121 CACATGCAATGATGATGATGTCGCGACGCTGCTGCGCTTCACTGATGAGGAGACCTC 1180
QY 841 TGCCTGGGGTGAATGTCGAGAGCCCAATGATGATGTCGCGCGACACCTGAGTT 900
DB 1181 TGCCTGGGGTGAATGTCGAGAGCCCAATGATGATGTCGCGCGACACCTGAGTT 1240
QY 901 GTCCAGACCTGCAATGCTCTCATCTACCTTGTGATCCATCTGAGACACCGCATGG 960
DB 1241 GTCCAGACCTGCAATGCTCTCATCTACCTTGTGATCCATCTGAGACACCGCATGG 1300
QY 961 GACATGATGATGATGATGTCGCGACGCGCTACATGATCTTGGGTACGCGATGCTGC 1020
DB 1301 GACATGATGATGATGATGTCGCGACGCGCTACATGATCTTGGGTACGCGATGCTGC 1360
QY 1021 CCCGAGGTCAATTAATGACATCATTAAGCGGAGCTCATTTGGGAGCTCATGTTCCGCTTGG 1080
DB 1361 CCCGAGGTCAATTAATGACATCATTAAGCGGAGCTCATTTGGGAGCTCATGTTCCGCTTGG 1420
QY 1081 TACTTCTGATGACGAGAGCGTGGGAGAAAGTGTGTCAATCTTGTGTCGCGCGG 1140
DB 1421 TACTTCTGATGACGAGAGCGTGGGAGAAAGTGTGTCAATCTTGTGTCGCGCGG 1480
QY 1141 GTGAGCGGCGACCCCATCTGTTGGGGGTTCTGCGCGGAGAACACCGGAGCGCTCAC 1200
DB 1481 GTGAGCGGCGACCCCATCTGTTGGGGGTTCTGCGCGGAGAACACCGGAGCGCTCAC 1540
QY 1201 AGCTTAATTTGACATGAGCGCCCGACGAGAAATCAGCTCTTAAACCAATGAGAGCTGG 1260
DB 1541 AGCTTAATTTGACATGAGCGCCCGACGAGAAATCAGCTCTTAAACCAATGAGAGCTGG 1600
QY 1261 CACATCAACCGCACCGGCTTGAATCTGCAATGATCTCTTGCACACCGGCTTTATCGGCT 1320
DB 1601 CACATCAACCGCACCGGCTTGAATCTGCAATGATCTCTTGCACACCGGCTTTATCGGCT 1660
QY 1321 CTGTTCTACACCCACAGCTTCACTGTCAGAGATGTCGGAAGCATGCGCTGCGCG 1380
DB 1661 CTGTTCTACACCCACAGCTTCACTGTCAGAGATGTCGGAAGCATGCGCTGCGCG 1720
QY 1381 AGTATCGAGGCTTCCGGGTGGGATGGGCGCTTGCATATGAGATTAATGTCACCAAT 1440
DB 1721 AGTATCGAGGCTTCCGGGTGGGATGGGCGCTTGCATATGAGATTAATGTCACCAAT 1780
QY 1441 CCAAGAGATGAGACCTTAATGCTGTGACCTACCCACCAAGGCAATGTGCGTGTCTCC 1500
DB 1781 CCAAGAGATGAGACCTTAATGCTGTGACCTACCCACCAAGGCAATGTGCGTGTCTCC 1840
QY 1501 GCGAAGCTGTGTGCGCCAGTGTACTGTTTACACCCCGGCGGAGGAGTGTGAGGAGC 1560
DB 1841 GCGAAGCTGTGTGCGCCAGTGTACTGTTTACACCCCGGCGGAGGAGTGTGAGGAGC 1900
QY 1561 ACCGACAGGCTTGGAGCGCCACTTACACGTCGGGAGGAGATGACAGATGTCCTCTA 1620
DB 1901 ACCGACAGGCTTGGAGCGCCACTTACACGTCGGGAGGAGATGACAGATGTCCTCTA 1960
QY 1621 TTGAACAGCACTGACCAACGCTGGGGTCAATGTTGGCTGACAGTGAATGAACTCTTCT 1680
DB 1961 TTGAACAGCACTGACCAACGCTGGGGTCAATGTTGGCTGACAGTGAATGAACTCTTCT 2020
QY 1681 GGCTACACCAAGACTTGGCGGACCAACCTGCGATCTAGAGCTGACTTCAACGCGAGC 1740
DB 2021 GGCTACACCAAGACTTGGCGGACCAACCTGCGATCTAGAGCTGACTTCAACGCGAGC 2080
QY 1741 ACGGACCTGTGTGCGCCACGAGCTGTTTGAAGAGCATCTGATACCTTACCTTCAAA 1800

D	b	2081	ACGGA	CTGTTGTG	CCCA	CGA	CTGTTT	TGGA	GCAT	CTGAT	TAC	CA	CTTAC	CT	CA	2140																
O	y	1801	TGCG	CTCTG	GGCC	CTG	CTCAG	CG	CCAA	GGTGC	CTGAT	GCAT	ACCC	CTA	CAG	CTCTG	1860															
D	b	2141	TGCG	CTCTG	GGCC	CTG	CTCAG	CG	CCAA	GGTGC	CTGAT	GCAT	ACCC	CTA	CAG	CTCTG	2200															
O	y	1661	CATT	ACCC	CTG	CA	CAG	TTA	CTAT	TAC	CA	CTT	CAAA	ATA	AGGA	TGTA	TG	TG	GGAG	GGGG	CTT	1920										
D	b	2201	CATT	ACCC	CTG	CA	CAG	TTA	CTAT	TAC	CA	CTT	CAAA	ATA	AGGA	TGTA	TG	TG	GGAG	GGGG	CTT	2260										
O	y	1921	GAG	CA	CAG	CT	CA	CGG	CTG	CA	TG	CA	ATT	CA	CT	CGT	GGG	GA	T	CGT	TG	CA	CTT	TG	GA	GA	C	1980				
D	b	2261	GAG	CA	CAG	CT	CA	CGG	CTG	CA	TG	CA	ATT	CA	CT	CGT	GGG	GA	T	CGT	TG	CA	CTT	TG	GA	GA	C	2320				
O	y	1981	AGAG	CA	CAG	AA	GT	CA	AGT	CT	CT	CTT	TG	TG	GA	CT	CA	CA	CG	GA	AT	TG	GG	CA	ATT	T	CA	CT	2040			
D	b	2321	AGAG	CA	CAG	AA	GT	CA	AGT	CT	CT	CTT	TG	TG	GA	CT	CA	CA	CG	GA	AT	TG	GG	CA	ATT	T	CA	CT	2380			
O	y	2041	TGCT	TTT	ACT	GG	CA	CTG	CC	CG	CTT	GT	CG	A	CTG	GA	CT	GGT	CTT	CT	CA	CA	CT	CA	CA	AA	A	CA	T	2100		
D	b	2381	TGCT	TTT	ACT	GG	CA	CTG	CC	CG	CTT	GT	CG	A	CTG	GA	CT	GGT	CTT	CT	CA	CA	CT	CA	CA	AA	A	CA	T	2440		
O	y	2101	GTGA	CGT	CA	ATT	CA	TG	TAT	ATG	CG	CT	AT	CA	CT	TG	CC	CT	CA	CA	AA	TA	CA	T	GT	CC	GA	TG	2160			
D	b	2441	GTGA	CGT	CA	ATT	CA	TG	TAT	ATG	CG	CT	AT	CA	CT	TG	CC	CT	CA	CA	AA	TA	CA	T	GT	CC	GA	TG	2500			
O	y	2161	GAGT	GG	GA	T	A	CT	CT	T	A	T	T	C	T	G	T	C	T	T	A	C	G	G	A	G	C	T	G	T	2220	
D	b	2501	GAGT	GG	GA	T	A	CT	CT	T	A	T	T	C	T	G	T	C	T	T	A	C	G	G	A	G	C	T	G	T	2560	
O	y	2221	ATG	CT	CA	T	CTT	G	T	G	GG	C	A	G	G	C	C	A	G	A	C	A	G	A	C	T	GA	A	G	A	G	2280
D	b	2561	ATG	CT	CA	T	CTT	G	T	G	GG	C	A	G	G	C	C	A	G	A	C	A	G	A	C	T	GA	A	G	A	G	2620
O	y	2281	GCG	AG	CG	AG	CT	AG	TG	GA	A	TG	CGT	C	CT	A	T	A	T	T	G	C	A	T	T	T	T	T	T	T	T	2340
D	b	2621	GCG	AG	CG	AG	CT	AG	TG	GA	A	TG	CGT	C	CT	A	T	A	T	T	G	C	A	T	T	T	T	T	T	T	T	2680
O	y	2341	TAC	AT	CA	A	G	G	T	C	G	G	T	A	G	T	C	C	C	C	T	A	G	T	C	C	T	A	T	C	C	2400
D	b	2681	TAC	AT	CA	A	G	G	T	C	G	G	T	A	G	T	C	C	C	C	T	A	G	T	C	C	T	A	T	C	C	2740
O	y	2401	AGC	T	ACT	AG	CT	CT	A	G	A	T	TG	C	C	CA	A	G	G	CT	A	T	AT	G	A	T	G	A	C	A	T	2460
D	b	2741	AGC	T	ACT	AG	CT	CT	A	G	A	T	TG	C	C	CA	A	G	G	CT	A	T	AT	G	A	T	G	A	C	A	T	2800
O	y	2461	C																													

Db	3161	GCGCTACTAGCCCTTGGCAGGTTGGACTGGACCTTAATCTATAGACACACTCACCCCTATG	3220
Qy	2881	TCGGATTBGGCTGCTAGTGGCTCTGCGGAGACTGGCGGTGGCTTGTAGCCTATCATTTTC	2940
Db	3221	TCGGATTBGGCTGCTAGTGGCTCTGCGGAGACTGGCGGTGGCTTGTAGCCTATCATTTTC	3280
Qy	2941	AGTCCGATBGAAGAAAGTCAATTGCTGGGGAGACGGACAGCTGCTTGTGGGGAATT	3000
Db	3281	AGTCCGATBGAAGAAAGTCAATTGCTGGGGAGACGGACAGCTGCTTGTGGGGAATT	3340
Qy	3001	TTAACCGGACTTCCCGTGTCCGCCGACTTGTGGAGGTCTCTTGGCCAGCTGAT	3060
Db	3341	TTAACCGGACTTCCCGTGTCCGCCGACTTGTGGAGGTCTCTTGGCCAGCTGAT	3400
Qy	3061	GGCTATACCTCCAAAGGGGTGGAGTCTTGTGCCCCCACTGACTGCTTAAGCCGACGACACA	3120
Db	3401	GGCTATACCTCCAAAGGGGTGGAGTCTTGTGCCCCCACTGACTGCTTAAGCCGACGACACA	3460
Qy	3121	CGTGGCCCTTTTGGGCAACCATAGTGGTAGCATGACGGGGCGCGACAAAGACAAGACT	3180
Db	3461	CGTGGCCCTTTTGGGCAACCATAGTGGTAGCATGACGGGGCGCGACAAAGACAAGACT	3520
Qy	3181	GGGGAAATTCAAGGTCTGTCCACAGTCACTAGTCTTTCTCGGAAATCATCTTCGGGG	3240
Db	3521	GGGGAAATTCAAGGTCTGTCCACAGTCACTAGTCTTTCTCGGAAATCATCTTCGGGG	3580
Qy	3241	GTTTGTGTGGACTGTCTACATGAGAGTGGCAACAAGACTCTGGCGGGCTCACGGGGTCCG	3300
Db	3581	GTTTGTGTGGACTGTCTACATGAGAGTGGCAACAAGACTCTGGCGGGCTCACGGGGTCCG	3640
Qy	3301	GTCACGCGAGATGTATCTCCAGTGTCTGAGGGGGACTTAAGAGGTGGCCGAGCCCTTGGG	3360
Db	3641	GTCACGCGAGATGTATCTCCAGTGTCTGAGGGGGACTTAAGAGGTGGCCGAGCCCTTGGG	3700
Qy	3361	ACTAAATTTTGGAGCGGTGACGATGTGGAGCGGTGCACCTGTACCTGTGACGCGGAAAC	3420
Db	3701	ACTAAATTTTGGAGCGGTGACGATGTGGAGCGGTGCACCTGTACCTGTGACGCGGAAAC	3760
Qy	3421	GCTGATGTCACTCCGGCTCGAAGAAGCGGGGACAAACGGGGAGGCTACTTCCCGAGA	3480
Db	3761	GCTGATGTCACTCCGGCTCGAAGAAGCGGGGACAAACGGGGAGGCTACTTCCCGAGA	3820
Qy	3481	CCTCTTTTCACCTTGAAGGGGTCTCTCAGAGAGCCCGGTGTGCTATGCCCCAGGGGCAAGCT	3540
Db	3821	CCTCTTTTCACCTTGAAGGGGTCTCTCAGAGAGCCCGGTGTGCTATGCCCCAGGGGCAAGCT	3880
Qy	3541	GTCGGAGCTTCCGGGAGAGCTGTGTGCTCTGCGGGGCTGTGCTAAGTCCATAGATTTGATC	3600
Db	3881	GTCGGAGCTTCCGGGAGAGCTGTGTGCTCTGCGGGGCTGTGCTAAGTCCATAGATTTGATC	3940
Qy	3601	CCCGTTGAGACCTTGACATCTGTACCGCGGTCCCCACCTTATGTGACAAACAGCACCA	3660
Db	3941	CCCGTTGAGACCTTGACATCTGTACCGCGGTCCCCACCTTATGTGACAAACAGCACCA	4000
Qy	3661	CCTGTGTGGCCCAAGACCTTACAGGTGGGGATCTTGTGATGGCCCGGACCTGGCAGTGGAAAG	3720
Db	4001	CCTGTGTGGCCCAAGACCTTACAGGTGGGGATCTTGTGATGGCCCGGACCTGGCAGTGGAAAG	4060
Qy	3721	AGCACCAAGTCTGTGTGACATGTCTGACGGGGATATAAAGTGTAGTGTCTTAATCCC	3780
Db	4061	AGCACCAAGTCTGTGTGACATGTCTGACGGGGATATAAAGTGTAGTGTCTTAATCCC	4120
Qy	3781	TCAGTGGCTGCAACCTTGGGGTTTGGGGGCTATCTTGTCTAAGGACATGGCATTAATCCC	3840
Db	4121	TCAGTGGCTGCAACCTTGGGGTTTGGGGGCTATCTTGTCTAAGGACATGGCATTAATCCC	4180
Qy	3841	AACATTAGAGCTGAGTCAAGACTGTGACGACCGGGGCGCCATCAAGTACTCCACATAT	3900
Db	4181	AACATTAGAGCTGAGTCAAGACTGTGACGACCGGGGCGCCATCAAGTACTCCACATAT	4240
Qy	3901	GCGAAATTCCTCGCCGATGGGGGCTGTGCGGGCGCGCTTACGACATCATATGTGAT	3960
Db	4241	GCGAAATTCCTCGCCGATGGGGGCTGTGCGGGCGCGCTTACGACATCATATGTGAT	4300

QY 3961 GAATGCCATGCGGTGAGCTTACCAACATCTTGGCATGGGAACAGTCTTGATCAAGCA 4020
DB 4301 GAATGCCATGCGGTGAGCTTACCAACATCTTGGCATGGGAACAGTCTTGATCAAGCA 4360
QY 4021 GAGACAGCTGGGGTTCAGACTTAAGTCTGCTGCTAAGCTAAGCCCTTGGGTCAAGTACA 4080
DB 4361 GAGACAGCTGGGGTTCAGACTTAAGTCTGCTGCTAAGCTAAGCCCTTGGGTCAAGTACA 4420
QY 4081 ACCCCCCCAACCATAGAGAGGTGGCCCTTGGGAGAGAGGGCCAGATCCCTTCTAT 4140
DB 4421 ACCCCCCCAACCATAGAGAGGTGGCCCTTGGGAGAGAGGGCCAGATCCCTTCTAT 4480
QY 4141 GGGAGGGCGATTCCTCCCTGCTTATCATCAAGGGAGGAAGATCTTGATCTTCCGATTTCA 4200
DB 4481 GGGAGGGCGATTCCTCCCTGCTTATCAATCAAGGGAGGAAGATCTTGATCTTCCGATTTCA 4540
QY 4201 AAGAAAAAGTGTACGAGCTCGCGCGGCCCTTCCGGGTATGGGCTTGAATCAAGTGGCA 4260
DB 4541 AAGAAAAAGTGTACGAGCTCGCGCGGCCCTTCCGGGTATGGGCTTGAATCAAGTGGCA 4600
QY 4261 TACTACAGAGGGTGGAGCTCTCCGTATATCAACTCAAGGAGACGTAGTGTGTCGCC 4320
DB 4601 TACTACAGAGGGTGGAGCTCTCCGTATATCAACTCAAGGAGACGTAGTGTGTCGCC 4660
QY 4321 ACCGAGGCCCTCATNGACGGGTATACGGGGACTTTGACTCCGGATCGACTGCAAGTGA 4380
DB 4661 ACCGAGGCCCTCATNGACGGGTATACGGGGACTTTGACTCCGGATCGACTGCAAGTGA 4720
QY 4381 GCGGTCACTCAAGTTGTAGACTTCAAGTTAGACCCCATTTCAACATACCAACAGATT 4440
DB 4721 GCGGTCACTCAAGTTGTAGACTTCAAGTTAGACCCCATTTCAACATACCAACAGATT 4780
QY 4441 GTCCCTCAAGACCTGTCTCAAGTACGAGCGCGGGGTGCAAGGGTATGGGAGAGACTG 4500
DB 4781 GTCCCTCAAGACCTGTCTCAAGTACGAGCGCGGGGTGCAAGGGTATGGGAGAGACTG 4840
QY 4501 GGGATTTATAGATATGTTTCCACTGAGAGAGAGCCCTCAAGAAATGTTTGAACAGTATG 4560
DB 4841 GGGATTTATAGATATGTTTCCACTGAGAGAGAGCCCTCAAGAAATGTTTGAACAGTATG 4900
QY 4561 CTCTGTAGAGTCTACGACGAGAGGGCGCATGGTATAGCTCAACCATCGAGACCAACC 4620
DB 4901 CTCTGTAGAGTCTACGACGAGAGGGCGCATGGTATAGCTCAACCATCGAGACCAACC 4960
QY 4621 GTCAAGCTCAGGGCGTATTTCAACAGCGCGGTTTGTGTGTGCCAAGCAATCTTGAG 4680
DB 4961 GTCAAGCTCAGGGCGTATTTCAACAGCGCGGTTTGTGTGTGCCAAGCAATCTTGAG 5020
QY 4681 TTTTGGAGAGGCAATTTTCAACCGGCTTCAACATAGATGCCACTTCTTCCCAACA 4740
DB 5021 TTTTGGAGAGGCAATTTTCAACCGGCTTCAACATAGATGCCACTTCTTCCCAACA 5080
QY 4741 AAGCAATCGGGGAAAAATTTGCAATCTTAACAGCTTACAGGCTACAGGCTACAGTGTG 4800
DB 5081 AAGCAATCGGGGAAAAATTTGCAATCTTAACAGGCTTACAGGCTACAGTGTGCGTAGG 5140
QY 4801 GCCAAAGCCCCCCCCCTGCTGAGAGCTATGTGAGTGTGACTGCACTCAAGCCC 4860
DB 5141 GCCAAAGCCCCCCCCCTGCTGAGAGCTATGTGAGTGTGACTGCACTCAAGCCC 5200
QY 4861 ACACTGTGGGCCCCCAACCTCTCTGTATCCGCTGTGTATCAACAGAGGTAC 4920
DB 5201 ACACTGTGGGCCCCCAACCTCTCTGTATCCGCTGTGTATCAACAGAGGTAC 5260
QY 4921 CTCAACATCCCGTGAAGAAATATGAGCACTGACCTGACAGCCGACTTGAAGTATG 4980
DB 5261 CTCAACATCCCGTGAAGAAATATGAGCACTGACCTGACAGCCGACTTGAAGTATG 5320
QY 4981 ACCAGACATGGGTCTTTGGCAGGGGAGTCTTTGGCGGCGCTCGCGCGTATTTGCGCG 5040
DB 5321 ACCAGACATGGGTCTTTGGCAGGGGAGTCTTTGGCGGCGCTCGCGCGTATTTGCGCG 5380

QY 5041 ACCGGGTGTGTTGATCATCGGCGCTTGSCACATTAACGAGGCGGCTGTGGCGG 5100
DB 5381 ACCGGGTGTGTTGATCATCGGCGCTTGSCACATTAACGAGGCGGCTGTGGCGG 5440
QY 5101 GACAAAGAGTCTCTATAGAGCTTTTATGATGATGAGGAATGTGCTTGAAGGCGGCT 5160
DB 5441 GACAAAGAGTCTCTATAGAGCTTTTATGATGATGAGGAATGTGCTTGAAGGCGGCT 5500
QY 5161 CTGATTGAAGAGGGGAGCGGATAGCGGAGATGCTGAAGTCCAAAGTCCAAAGCTTATG 5220
DB 5501 CTGATTGAAGAGGGGAGCGGATAGCGGAGATGCTGAAGTCCAAAGTCCAAAGCTTATG 5560
QY 5221 CAGCAAGCTTCCAAAGGCTCAAGACATCAACCACTGTGAGGCTTACGAGCCCAAG 5280
DB 5561 CAGCAAGCTTCCAAAGGCTCAAGACATCAACCACTGTGAGGCTTACGAGCCCAAG 5620
QY 5281 GTAGAACATTTCTGGGCCAAACATGTGGAACTTATAGCGGATCCAAATCTTGCA 5340
DB 5621 GTAGAACATTTCTGGGCCAAACATGTGGAACTTATAGCGGATCCAAATCTTGCA 5680
QY 5341 GAGCTATCAACACTGCGCAGGGAACCTGCAAGTACCTTCAATGAGGCTTCAAGTCCGC 5400
DB 5681 GAGCTATCAACACTGCGCAGGGAACCTGCAAGTACCTTCAATGAGGCTTCAAGTCCGC 5740
QY 5401 CTCACAGTCCGCTGTCAACAGACCACTATCTTCTCAACATTTTGGGGGCTGGCTA 5460
DB 5741 CTCACAGTCCGCTGTCAACAGACCACTATCTTCTCAACATTTTGGGGGCTGGCTA 5800
QY 5461 GCATCCCAATTTGACACACCCCGGGGGCCACTGGCTTGTGTCAAGTGGCTATGGGA 5520
DB 5801 GCATCCCAATTTGACACACCCCGGGGGCCACTGGCTTGTGTCAAGTGGCTATGGGA 5860
QY 5521 GCTCCGTATAGGAGCTATAGGCTTATAGTATAGTGTAGTGAATCTTGGAGAGGTATGT 5580
DB 5861 GCTCCGTATAGGAGCTATAGGCTTATAGTATAGTGTAGTGAATCTTGGAGAGGTATGT 5920
QY 5581 GCGGGCAATTTGGGGGCTCTCGTGCATTCAGATCATGCTGGCGAAGCCCTCATG 5640
DB 5921 GCGGGCAATTTGGGGGCTCTCGTGCATTCAGATCATGCTGGCGAAGCCCTCATG 5980
QY 5641 GAGATGTGCTCAACTGTGCTGCTGGAATTTCTCTCCGGGTGCTTGTATGGAGTCT 5700
DB 5981 GAGATGTGCTCAACTGTGCTGCTGGAATTTCTCTCCGGGTGCTTGTATGGAGTCT 6040
QY 5701 ATCTGCGCGGCAATTTGCGCGGACACGATGGAACCGGGGAAAGCGCTTCAATGATG 5760
DB 6041 ATCTGCGCGGCAATTTGCGCGGACACGATGGAACCGGGGAAAGCGCTTCAATGATG 6100
QY 5761 AATGACTCATTTGCTTGTCCAGAGGAATCAAGTGGCCGCCCAACCACTAGTAGC 5820
DB 6101 AATGACTCATTTGCTTGTCCAGAGGAATCAAGTGGCCGCCCAACCACTAGTAGC 6160
QY 5821 GAGTGGATGCGTGCAGAGCGTGTACCAACTATCTTGCTCTCTTACATTAACAGCTG 5880
DB 6161 GAGTGGATGCGTGCAGAGCGTGTACCAACTATCTTGCTCTCTTACATTAACAGCTG 6220
QY 5881 CTGAGAAAGTCCAACTGGAATTAAGTGAAGATGCGCCCATCCCATGCGGCGCTGTG 5940
DB 6221 CTGAGAAAGTCCAACTGGAATTAAGTGAAGATGCGCCCATCCCATGCGGCGCTGTG 6280
QY 5941 CTCGCGAGTGTGGGACTGGGTTTGACCACTCTTAACAGACTTTAAATAATTTGGCTGAC 6000
DB 6281 CTCGCGAGTGTGGGACTGGGTTTGACCACTCTTAACAGACTTTAAATAATTTGGCTGAC 6340
QY 6001 TCCAAATTAATTTCCCAAGATGCGGCGCTTCCCTTGTCTCTGTCAAAAAGGCTACAG 6060
DB 6341 TCCAAATTAATTTCCCAAGATGCGGCGCTTCCCTTGTCTCTGTCAAAAAGGCTACAG 6400
QY 6061 GGGGTGTGGGCGGCACTGAGTATGACCAACAGGATGCTTGGCGGCGCAATATCTCT 6120
DB 6401 GGGGTGTGGGCGGCACTGAGTATGACCAACAGGATGCTTGGCGGCGCAATATCTCT 6460
QY 6121 GGCATGTCCGCTTGTGGCTTCATGAGATCAAGGGGCTTAAGACTGATGAATATCTGG 6180

821 GGGGTTAAATTTTGCAACAGGGAATTCACCCGGTGTCTCTTTTCTATCTTCTGCGCC 880
QY CTGCTGCTCTGCATCAACCAACCCGGTCTCCGCTGCGAAGTGAAGAACATCAGTACCGAC 600
DB CTGCTGCTCTGCATCAACCAACCCGGTCTCCGCTGCGAAGTGAAGAACATCAGTACCGAC 940
QY TACATGTGTACTAAGACTGACCAATGACATGACATTAACCTGGCAGCTCCAGGCTGTCTTC 660
DB TACATGTGTACTAAGACTGACCAATGACATGACATTAACCTGGCAGCTCCAGGCTGTCTTC 1000
QY CTCACGTCCTCCGGGGTGGTCCCGTGGGAGAAAGTGGGAAATGACATCTCAGTGTGATTA 720
DB CTCACGTCCTCCGGGGTGGTCCCGTGGGAGAAAGTGGGAAATGACATCTCAGTGTGATTA 1060
QY CCGGTCTCAACCAATGTGGCCGTGACAGCGCCCGCGCTCAGCAGAGGCTTACGAGACG 780
DB CCGGTCTCAACCAATGTGGCCGTGACAGCGCCCGCGCTCAGCAGAGGCTTACGAGACG 1120
QY CACATCGACATGTGTGTGATGTCCGCCACGCTCTGCTCTGCCCTCTACGTGGGAGACCTC 840
DB CACATCGACATGTGTGTGATGTCCGCCACGCTCTGCTCTGCCCTCTACGTGGGAGACCTC 1180
QY TGGCGGTGGGTGTATGCTGCGACAGCCCAATGTTGATTTGTCTGGCCGACAGCAACATGTTT 900
DB TGGCGGTGGGTGTATGCTGCGACAGCCCAATGTTGATTTGTCTGGCCGACAGCAACATGTTT 1240
QY GTCAAGAAGTGCATTTGCTCATCTTACCCTGTGATCACTGACACACGATGACATG 960
DB GTCAAGAAGTGCATTTGCTCATCTTACCCTGTGATCACTGACACACGATGACATG 1300
QY GACATGATGATGAATCTGTGCGCCACCGCTACCATGATCTTGGCGTACCGCATGTGCTTC 1020
DB GACATGATGATGAATCTGTGCGCCACCGCTACCATGATCTTGGCGTACCGCATGTGCTTC 1360
QY GACATGATGATGAATCTGTGCGCCACCGCTACCATGATCTTGGCGTACCGCATGTGCTTC 1080
DB GACATGATGATGAATCTGTGCGCCACCGCTACCATGATCTTGGCGTACCGCATGTGCTTC 1420
QY CCGGAGTCTATATGACATCACTTGGCGGCTCATTTGGGCGTCAATGCTTACCC 1040
DB CCGGAGTCTATATGACATCACTTGGCGGCTCATTTGGGCGTCAATGCTTACCC 1480
QY TACTTCTCTATGACAGGAGCGTGGGCGAAAGTCTTGTGATCTTCTGTTGACCGCGCG 1140
DB TACTTCTCTATGACAGGAGCGTGGGCGAAAGTCTTGTGATCTTCTGTTGACCGCGCG 1480
QY GTGAGCGCGCGACCCATCTGTGTTGGGGTCTTCCCGCGACAGCAACCGGCGCTCAC 1200
DB GTGAGCGCGCGACCCATCTGTGTTGGGGTCTTCCCGCGACAGCAACCGGCGCTCAC 1540
QY GTGAGCGCGCGACCCATCTGTGTTGGGGTCTTCCCGCGACAGCAACCGGCGCTCAC 1260
DB GTGAGCGCGCGACCCATCTGTGTTGGGGTCTTCCCGCGACAGCAACCGGCGCTCAC 1600
QY GGCATGTTCTCTGTTGTGTCAGGAGCAAAATCCAGCTCATCAACCAATGAGCAAGTTGG 1320
DB GGCATGTTCTCTGTTGTGTCAGGAGCAAAATCCAGCTCATCAACCAATGAGCAAGTTGG 1660
QY CACATCAACCGACCGCCCTGAATGCAATGACTCTTTGCAACACGGCTTCTCTGCGTCA 1380
DB CACATCAACCGACCGCCCTGAATGCAATGACTCTTTGCAACACGGCTTCTCTGCGTCA 1720
QY CTGTTCTACACCCACAGCTTCAACTGTCAGGATGTCCGGAACGATGTCCGCTGCGGC 1440
DB CTGTTCTACACCCACAGCTTCAACTGTCAGGATGTCCGGAACGATGTCCGCTGCGGC 1780
QY AGTATCGAGGCGCTTCCGGGTGGGATGGGGCGCTTGCAATAGAGGATTAATGCAAT 1500
DB AGTATCGAGGCGCTTCCGGGTGGGATGGGGCGCTTGCAATAGAGGATTAATGCAAT 1840
QY CCGAGAGATATGAGACCGTATGTGCGCACTAACCAAGACAGTGTGTATCTTCC 1560
DB CCGAGAGATATGAGACCGTATGTGCGCACTAACCAAGACAGTGTGTATCTTCC 1900
QY GCGAGAGATATGAGACCGTATGTGCGCACTAACCAAGACAGTGTGTATCTTCC 1620
DB GCGAGAGATATGAGACCGTATGTGCGCACTAACCAAGACAGTGTGTATCTTCC 1960
QY ACCGATATGACTTGGAGCGCCACTTACAGGTGGGGAGAAATGAGACAGATGTCTTCTTA 1680

QY TTGAAACGACCTCGACACCGCTGGGTCATGTTGGCTGCAAGTGAATGAACTTCT 1620
DB TTGAAACGACCTCGACACCGCGAGGGTCAATGTTGGCTGCAAGTGAATGAACTTCT 2020
QY GGTCAACCAAGACTTGGGCGGACACCGTCCGATCTAGAGCTGACTTCAACGCGAC 1740
DB GGTCAACCAAGACTTGGGCGGACACCGTCCGATCTAGAGCTGACTTCAACGCGAC 2080
QY ACCGACTGTGTGCGCCACAGACTGTTTAAAGAACATCTGATACCACTTACCTGAA 1800
DB ACCGACTGTGTGCGCCACAGACTGTTTAAAGAACATCTGATACCACTTACCTGAA 2140
QY TGGGCTCTGGGCGCTGAGCTCAGCGCAAGGTGCTGATGACTTACCTTACAGGCTTGG 1860
DB TGGGCTCTGGGCGCTGAGCTCAGCGCAAGGTGCTGATGACTTACCTTACAGGCTTGG 2200
QY CATTAACCTGCGACACTTAACTATACCACTTCAAAATGAATGATATGGGAGGGT 1920
DB CATTAACCTGCGACACTTAACTATACCACTTCAAAATGAATGATATGGGAGGGT 2260
QY GAGCAGAGGCTCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
DB GAGCAGAGGCTCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2320
QY AGAGACGAAGTCACTGCTCTCTTGTGACCTCAGCAGGAGTGGGCAATTTTACCT 2040
DB AGAGACGAAGTCACTGCTCTCTTGTGACCTCAGCAGGAGTGGGCAATTTTACCT 2380
QY TGTCTTACTCGGACCTGCGCGCTTGTGATGATGATGATGATGATGATGATGATGATGAT 2100
DB TGTCTTACTCGGACCTGCGCGCTTGTGATGATGATGATGATGATGATGATGATGATGAT 2440
QY GTGAGAGTAACTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
DB GTGAGAGTAACTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2500
QY GAGTGGGTAGTACTTATTTCTGCTTGTAGCGGAGCGGAGTGGGCTTGTATG 2220
DB GAGTGGGTAGTACTTATTTCTGCTTGTAGCGGAGCGGAGTGGGCTTGTATG 2560
QY ATGCTCATCTGTTGGGCGACCGGCGAAGCACTAGAGAGTGGTCTTGTGACGCT 2280
DB ATGCTCATCTGTTGGGCGACCGGCGAAGCACTAGAGAGTGGTCTTGTGACGCT 2620
QY GCGAGCGAGCTGACTGCAATGAGCTTCTATATTTGTATCTTGTGCTGCTGCTG 2340
DB GCGAGCGAGCTGACTGCAATGAGCTTCTATATTTGTATCTTGTGCTGCTGCTG 2680
QY TACATCAAGGCTGGGTAGTCCCTTACCTATTTCCCTCACTGGCTGTGTCTT 2400
DB TACATCAAGGCTGGGTAGTCCCTTACCTATTTCCCTCACTGGCTGTGTCTT 2740
QY AGCTTACTGCTCTAGCAATTTGCCCAACAGGCTTATGATGATGATGATGATGATGAT 2460
DB AGCTTACTGCTCTAGCAATTTGCCCAACAGGCTTATGATGATGATGATGATGATGAT 2800
QY CAGATAGAGAGCTGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
DB CAGATAGAGAGCTGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2860
QY ACCCTTCTAGCGGTTTTTGT 2580
DB ACCCTTCTAGCGGTTTTTGT 2920
QY GTCAAGAGTGGGCAACCACTATGCAAGTGGGCGGCGGATGAGATGATGATGATGATGAT 2640
DB GTCAAGAGTGGGCAACCACTATGCAAGTGGGCGGCGGATGAGATGATGATGATGATGAT 2980
QY GTGAGCATATTTTCAACCAAGT 2700
DB GTGAGCATATTTTCAACCAAGT 3040

QY 2701 GGGCTGCTTACCTCTTAAAGGTCTTTGACGCGCGTCCGTACTTCTGTCAGGGCTCAC 2760
DB |||||
QY 3041 GGGCTGCTTATCTCTAAAGGTCTTTGACGCGTGTCCGTACTTCTGTCAGGGCTCAC 3100
DB |||||
QY 2761 GCTCTACTGAGATGTGCACCATGCAAGGCACTCTGCGGGGGCAGGTACGTCCAGATG 2820
DB |||||
QY 3101 GCTCTACTGAGATGTGCACCATGTAAGGCATCTCCGGGGGGTAAGTAGTCAATG 3160
DB |||||
QY 2821 GCGCTACTGAGCCCTGGCAGGTGGACTGGCATTACATTAATGACCACTCAACCCCTATG 2880
DB |||||
QY 3161 GTGCTACTGAGCCCTGGCAGGTGGACTGGCATTACATTAATGACCACTCAACCCCTATG 3220
DB |||||
QY 2881 TCGGATTGGCTGTGTAATGCTGTGCGGGACTTGGCGGCTCCGTGGAGCTTAATCTTC 2940
DB |||||
QY 3221 TCGGATTGGCTGTGTAATGCTGTGCGGGACTTGGCGGCTCCGTGGAGCTTAATCTTC 3280
DB |||||
QY 2941 AGTCCGATGGAAGAAAGTCAATTTGTCTGGGAGCGGAGACAGCTGCTTGTGGGACAT 3000
DB |||||
QY 3281 AGTCCGATGGAAGAAAGTCAATTTGTCTGGGAGCGGAGACAGCTGCTTGTGGGATATC 3340
DB |||||
QY 3001 TTACACGAGCTTCCCGTGTCCGCCGACTTGTCTGGGAGGTCCTTCTGGCCCGAGCTGAT 3060
DB |||||
QY 3341 TTACACGAGCTTCCCGTGTCCGCCGACTTGTCTGGGAGGTCCTTCTGGCCCGAGCTGAT 3400
DB |||||
QY 3061 GGGCTATCTCTCAAGGGGTGAGTCTTCTGCGCCCATCACTGCTTACGCCCAAGCAGCA 3120
DB |||||
QY 3401 GGGCTATCTCTCAAGGGGTGAGTCTTCTGCGCCCATCACTGCTTATGCCCAGCAGACA 3460
DB |||||
QY 3121 CGTGGCCTTTTGGGACCATATGCTGTGAGCATAGACGCGGCGCGACAGACAGAGCCT 3180
DB |||||
QY 3461 CGGCGCTTTTGGGACCATATGCTGTGAGCATAGACGCGGCGCGACAGACAGAGCCT 3520
DB |||||
QY 3181 GGGGAAATTGAGGCTCTGTCCACAGTCACTAGTCTTCTCTGGGAAATCATCTCGGG 3240
DB |||||
QY 3521 GGGGAAATTGAGGCTCTGTCCACAGTCACTAGTCTTCTCTGGGAAATCATCTCGGG 3580
DB |||||
QY 3241 GTTTTGTGAGCTGTCTCACTGAGGCTGTGCAACAAGCTCTGCGCGGCTCAAGGGGTCCG 3300
DB |||||
QY 3581 GTCTTATGAGCTGTCTCACTGAGGCTGTGCAACAAGCTCTAGCGGCTCAAGGGGTCCG 3640
DB |||||
QY 3301 GTCAACGAGATGTATCTCACTGAGGCTGTGCAACAAGCTCTAGGGGTCCAGCGGCTG 3360
DB |||||
QY 3641 GTCAACGAGATGTATCTCACTGAGGCTGTGCAACAAGCTCTAGGGGTCCAGCGGCTG 3700
DB |||||
QY 3361 ACTTAATCTTTGAGAGCGGTGCACTGTGAGCGGTGCACTGTATCTGTGTCAAGCGGAC 3420
DB |||||
QY 3701 ACCAAATCTTTGAGAGCGGTGCACTGTGAGCGGTGCACTGTATCTGTGTCAAGCGGAC 3760
DB |||||
QY 3421 GCTGATGTCACTCCGCGCTCGAAGACGCGGGGACAAACGCGGAGGCTACTCTCCCGAGA 3480
DB |||||
QY 3761 GCTGATGTCACTCCGCGCTCGAAGACGCGGGGACAAACGCGGAGGCTACTCTCCCGAGA 3820
DB |||||
QY 3481 CCTCTTTTCACTTTGAAGGGGTCTCTGAGAGCCCGGTGCTATGCCCCAGGGGCAAGCT 3540
DB |||||
QY 3821 CCTCTTTTCACTTTGAAGGGGTCTCTGAGAGCCCGGTGCTATGCCCCAGGGGCAAGCT 3880
DB |||||
QY 3541 GTGAGAGTCTTCCGGGAGAGCTGTGTGCTCTCGGGGCTGTGAGTCAATGATTTTATC 3600
DB |||||
QY 3881 GTGAGAGTCTTCCGGGAGAGCTGTGTGCTCTCGGGGCTGTGAGTCAATGATTTTATC 3940
DB |||||
QY 3601 CCGGTTGAGACACTCGACATCTGTCAAGCGGTCTCCCGCACTTTAGTGAACAAGCAGCA 3660
DB |||||
QY 3941 CCGGTTGAGACACTCGACATCTGTCAAGCGGTCTCCCGCACTTTAGTGAACAAGCAGCA 4000
DB |||||
QY 3661 CCTGCTGTGCCAGACACTATCAGGTCTGGGTCTTTCATGCCCCGAGCTGCGAGTGAAG 3720
DB |||||
QY 4001 CCTGCTGTGCCAGACACTATCAGGTCTGGGTCTTTCATGCCCCGAGCTGCGAGTGAAG 4060
DB |||||
QY 3721 AGGACCAAGTCTCTGTGATATGCTCAAGGGGTATTAAGGCTAGTCTTAAATCCC 3780
DB |||||
QY 4061 AGGACCAAGTCTCTGTGATATGCTCAAGGGGTATTAAGGCTAGTCTTAAATCCC 4120
DB |||||
QY 3781 TCACTGTGCTGCAACCTCGGGGTTTGGGCGTACTTGTCTAAGGACATGCGCATCAATCCC 3840
DB |||||

DB 4121 TCGGTGCTGCAACCTCGGGGTTTGGGCGCTATCTGTCCAGGACATGCGCATCAATCCC 4180
QY 3841 AACATTAAGACTGAGAGTCAAGACTGTGACACCGGGGCGCCCATCACTACTCAATAT 3900
DB |||||
QY 4181 AACATTAAGACTGAGAGTCAAGACTGTGACACCGGGGCGCCCATCACTACTCAATAT 4240
DB |||||
QY 3901 GGGCAATTCCTCGCCGATGGGGGCTGTGGGGGCGGCGCTACAGCATCATATATGAT 3960
DB |||||
QY 4241 GGGCAATTCCTCGCCGATGGGGGCTGTGGGGGCGGCGCTACAGCATCATATATGAT 4300
DB |||||
QY 3961 GAATGCCATGCGGTGACTTACACCATCTTGGCATCGGAAACATCTTGTATCAAGCA 4020
DB |||||
QY 4301 GAATGCCATGCGGTGACTTACACCATCTTGGCATCGGAAACATCTTGTATCAAGCA 4360
DB |||||
QY 4021 GAGACAGCTGGGGTCAAGCTTACATCTGTGTGGCTTACAGCTACGCCCTGGGTCACTGACA 4080
DB |||||
QY 4361 GAGACAGCTGGGGTCAAGCTTACATCTGTGTGGCTTACAGCTACGCCCTGGGTCACTGACA 4420
DB |||||
QY 4081 ACCCCCCAGCCCAATPAGAGAGTGGGCGCTTGGGAGGAGGCGAGTCCCTTCTAT 4140
DB |||||
QY 4421 ACCCCCCAGCCCAATPAGAGAGTGGGCGCTTGGGAGGAGGCGAGTCCCTTCTAT 4480
DB |||||
QY 4141 GGGAGGCGATTCCTCTGTCTTACATCAAGGAGAAAGACATCTGATCTTCTGCAATCA 4200
DB |||||
QY 4481 GGGAGGCGATTCCTCTGTCTTACATCAAGGAGAAAGACATCTGATCTTCTGCAATCA 4540
DB |||||
QY 4201 AAGAAAAGTGTACAGAGCTTCGCGGCGGCGCTTGGGGTATGGGCTTGAATCACTAGTGA 4260
DB |||||
QY 4541 AAGAAAAGTGTACAGAGCTTCGCGGCGGCGCTTGGGGTATGGGCTTGAATCACTAGTGA 4600
DB |||||
QY 4261 TACTACAGAGGTTTGGAGCTCTCGTAAATCAACTCAGGAGACGTAGTGTGTGCC 4320
DB |||||
QY 4601 TACTACAGAGGTTTGGAGCTCTCGTAAATCAACTCAGGAGACGTAGTGTGTGCC 4660
DB |||||
QY 4321 ACCGAGCCCTCACTGACAGGGTATATCTGGGAACTTTGACTCGTGTATGCACTGCAAGTA 4380
DB |||||
QY 4661 ACCGAGCCCTCACTGACAGGGTATATCTGGGAACTTTGACTCGTGTATGCACTGCAAGTA 4720
DB |||||
QY 4381 GCGGTCACTCAAGTTTGAATTTCACTTGAACCCCATTCACATTAACCAACAGACT 4440
DB |||||
QY 4721 GCGGTCACTCAAGTTTGAATTTCACTTGAACCCCATTCACATTAACCAACAGACT 4780
DB |||||
QY 4441 GTCCCTCAAGAGCTGTCTCACTGAGCGAGCGCGGGGTGCGAGGCTAGGGGAAAGCTG 4500
DB |||||
QY 4781 GTCCCTCAAGAGCTGTCTCACTGAGCGAGCGCGGGGTGCGAGGCTAGGGGAAAGCTG 4840
DB |||||
QY 4501 GGCATTTATGATATGTTTCACTGAGTGAAGCGAGCTCAGGAAATGTTTGAACAGTATG 4560
DB |||||
QY 4841 GGCATTTATGATATGTTTCACTGAGTGAAGCGAGCTCAGGAAATGTTTGAACAGTATG 4900
DB |||||
QY 4561 CTCTGTGAGTGTACAGACGCGAGGGCGCATGATATGAGTCAACATTAAGAGACCA 4620
DB |||||
QY 4901 CTCTGTGAGTGTACAGACGCGAGGGCGCATGATATGAGTCAACATTAAGAGACCA 4680
DB |||||
QY 4621 GTGAGGCTCAAGGGGATTTTCAACAGCGCGGTTTGTGCTGTGCGCAAGACATCTTATG 4680
DB |||||
QY 4961 GTGAGGCTCAAGGGGATTTTCAACAGCGCGGTTTGTGCTGTGCGCAAGACATCTTATG 5020
DB |||||
QY 4681 TTTTGGAGAGGCACTTTTCAACCGGCTCAACATATGATGCCACTTCTTTCCCAACA 4740
DB |||||
QY 5021 TTTTGGAGAGGCACTTTTCAACCGGCTCAACATATGATGCCACTTCTTTCCCAACA 5080
DB |||||
QY 4741 AAGCAATCGGGGGAATTTTGCATCTTAAACGCTTACAGGCTACAGTGTGCGCTAGG 4800
DB |||||
QY 5081 AAGCAATCGGGGGAATTTTGCATCTTAAACGCTTACAGGCTACAGTGTGCGCTAGG 5140
DB |||||
QY 4801 GCCAAAGCCCCCCCCGCTCTGTGAGAGTCAATGGAAGTGTGTGACTCGACTCAAGCCC 4860
DB |||||
QY 5141 GCCAAAGCCCCCCCCGCTCTGTGAGAGTCAATGGAAGTGTGTGACTCGACTCAAGCCC 5200
DB |||||
QY 4861 ACACTGTGAGCCCAACCTCTCTGTACCGGTTGGGCTGTGTACCAAGAGGTACCC 4920
DB |||||

Dh 5201 AACTCGTGGGCCCAACCTCTCTCTGTAACCGTGTGGGCTGTGTAAACAGAGTACC 5260
Qy 4921 CTCAACATCCCGTGAAGAAATATACATGGCGACTGCAATGCAAGCGGACCTTTAGAGTCATG 4980
Dh 5261 CTACAGCACTCTGACGAAATACATGCCACCTGACAGCAAGCGACTTGAAGTATAG 5320
Qy 4981 ACCAGCAATGGGCTTTGGCAGAGGAGAGTCTTGGCGGCGCTGCGCGGATATTTGCTGGC 5040
Dh 5321 ACCAGCAATGGGCTTTAGTGGGAGGAGTCTTGGCGGCGCTGCGCGGATATTTGCTGGC 5380
Qy 5041 ACCGGGTGTGTGTGATGATGAGGCGGCTTGGCACTTAACAGCGAGCGCTGTGGCCCG 5100
Dh 5381 ACCGGGTGTGTGTGATGATGAGGCGGCTTGGCACTTAACAGCGAGCGCTGTGGCCCG 5440
Qy 5101 GACAAGAGGTCTCTATGAGGCTTTTGTATGATGAGATGAGAGATGTGCTTAGGCGGCT 5160
Dh 5441 GACAAGAGGTCTCTATGAGGCTTTTGTATGAGATGAGAGATGTGCTTAGGCGGCT 5500
Qy 5161 CTCAATTGAAGAGGCGAGCGGATAGCGGAGATGCTGAATGCTCAAGATCCAAAGCTTATG 5220
Dh 5501 CTCAATTGAAGAGGCGAGCGGATAGCGGAGATGCTGAATGCTCAAGATCCAAAGCTTATG 5560
Qy 5221 CAGCAAGCTTCCAAACAGGCTCAAGACATACACCCCTGTGACGGCTTACAGCGGAG 5280
Dh 5561 CAGCAAGCTTCCAAACAGGCTCAAGACATACACCCCTGTGACGGCTTCTTGGCCCAAG 5620
Qy 5281 GTAGAACAAATTCTGGGCGAAACACATGTGAACTTTAGCGGCAATCCAAATACCTGCA 5340
Dh 5621 GTAGAACAAATTCTGGGCGAAACACATGTGAACTTTAGCGGCAATCCAAATACCTGCA 5680
Qy 5341 GGAATATACACATGCGAGGGAACCTGTGAGTACCTGATGAGGCTTCAAGTGGCCGC 5400
Dh 5681 GGAATATACACATGCGAGGGAACCTGTGAGTACCTGATGAGGCTTCAAGTGGCCGC 5740
Qy 5401 CTCACCAATCCGCTGTCAACAAAGCACTATCTCTTCAACATTTTGGGGGCGCTGCTA 5460
Dh 5741 CTCACCAATCCGCTGTGTCACTAGCACATCTCTTCAACATTTTGGGGGCGCTGCTA 5800
Qy 5461 GCATCCCAATTTGACACACCGCGGGGCGCACTGGCTTGTGTGACGTGCGCTAGTGGGA 5520
Dh 5801 GCATCCCAATTTGCGGCTCTCCGCGGGGCGTACCGGCTTGTGTGACGTGCGCTAGTGGG 5860
Qy 5521 GCTGCGGTAGGCAATAGGCTTGAAGTGTCTAGTGAATCTCTGACAGGATAGT 5580
Dh 5861 GCTGCGGTAGGCAATAGGCTTGAAGTGTCTAGTGAATCTCTGACAGGATAGT 5920
Qy 5581 GCGGGCAATTTGGGGGCGCTGTGTGCAATTCAAAGATGATGCTGGGAGAAAGCCCTCATG 5640
Dh 5921 GCGGGCAATTTGGGGGCGCTGTGTGCAATTCAAAGATGATGCTGGGAGAAAGCCCTCATG 5980
Qy 5641 GAGATGTCTCAACCTTGTGCTGCTGAATCTGTCTCCGGGTGCTTGGTATGATGGAGTC 5700
Dh 5981 GAGATGTGTGTCAACCTTGTGCTGCTGAATCTGTCTCCGGGTGCTTGGTATGATGGAGTC 6040
Qy 5701 ATCTGCGGCGCATTTCTGCGCGACATGCGGAGCCGGGGAAGGCGCTGTCCAAATGAGATG 5760
Dh 6041 ATCTGCGGCGCATTTCTGCGCGACATGCGGAGCCGGGGAAGGCGCTGTCCAAATGAGATG 6100
Qy 5761 AATAGACTCATTTGCTTTGCTTCCAGAGAAATCAAGTGTGCGCCCAACCACTAGGTAGCG 5820
Dh 6101 AATAGACTCATTTGCTTTGCTTCCAGAGAAATCAAGTGTGCGCCCAACCACTAGGTAGCG 6160
Qy 5821 GAGTGCAGTGCCTGCGACAGGTGTGACCAACTATTTGCTTCCATTAACAGAGCTG 5880
Dh 6161 GAGTGCAGTGCCTGCGACAGGTGTGACCAACTATTTGCTTCCATTAACAGAGCTG 6220
Qy 5881 CTGAGAGACTTCACAACTGATTACTAGAGACTGCGCCATTCCTAGTGGCGGCTGTGG 5940
Dh 6221 CTGAGAGACTTCACAACTGATTACTAGAGACTGCGCCATTCCTAGTGGCGGCTGTGG 6280
Qy 5941 CTCCCGGAGTGTGGGACTGGGTTTGGCACTATCCCTTAACAGACTTTAAATTTGGTGAAC 6000
Dh 6281 CTCCCGGAGTGTGGGACTGGGTTTGGCACTATCCCTTAACAGACTTTAAATTTGGTGAAC 6340

Qy 6001 TCCAAATTAATCCCAAAGATGCCCGGCTCCCTTTGTCTCTGTCAAAAAGGGGTACAAG 6060
Dh 6341 TCCAAATTAATCCCAAAGATGCCCGGCTCCCTTTGTCTCTGTCAAAAAGGGGTACAAG 6400
Qy 6061 GCGGTGTGGGCGCGGCACTGGCATATGACCAACAGGTGTCTTTGGCGGCGCAATCTCT 6120
Dh 6401 GCGGTGTGGGCGCGGCACTGGCATATGACCAACAGGTGTCTTTGGCGGCGCGCAATCTCT 6460
Qy 6121 GCGAATGTCCGCTGGGCGCTCAATGAAATCAAGGGGCTTAAGACCTGCAATGAATCTGG 6180
Dh 6461 GCGAATGTCCGCTGGGCGCTCAATGAAATCAAGGGGCTTAAGACCTGCAATGAATCTGG 6520
Qy 6181 CAGGGAGCTTTCTATCAATTTGTTACAGAGAGGCACTGCTGTCGGAACCGCGGCA 6240
Dh 6521 CAGGGAGCTTTCTCCATCAATTTGTTACAGAGAGGCACTGCTGTCGGAACCGCGGCA 6580
Qy 6241 AACTTTAAGTGGCCATCTGAGAGGTGGGCGCTCAGAGTACGCGGAGGTGACGACGAC 6300
Dh 6581 AACTTTAAGTGGCCATCTGAGAGGTGGGCGCTCAGAGTACGCGGAGGTGACGACGAC 6640
Qy 6301 GGGTCAACCACTACATCAATGAGACTCACCACTGATTAACCTTGAAGTCCCTGCCAAT 6360
Dh 6641 GGGTCAACCACTACATCAATGAGACTCACCACTGATTAACCTTGAAGTCCCTGCCAAT 6700
Qy 6361 CCCTCTCCGAGTCTTTCTGGGTGACGAGAGTGCAGATCCATAGTTTCCGCCACA 6420
Dh 6701 CCCTCTCCGAGTCTTTCTGGGTGACGAGAGTGCAGATCCATAGTTTCCGCCACA 6760
Qy 6421 CCGAAGCGCTTTTCCGGGATGAGGTCTGCTGTGGGTGGGCTTAATTCATTTGTGCTG 6480
Dh 6761 CCGAAGCGCTTTTCCGGGATGAGGTCTGCTGTGGGTGGGCTTAATTCATTTGTGCTG 6820
Qy 6481 GGGTCCAGCTTCTTGGGCAACCTGAACCGGACACAGAGTATGATGCTCAATGCTTACA 6540
Dh 6821 GGGTCCAGCTTCTTGGGCAACCTGAACCGGAGTATGATGCTCAATGCTTACA 6880
Qy 6541 GATCATCTCATATCAACGCGAGAGACTGACGCGGCGTTTACGCGGGGGTCAACCCCA 6600
Dh 6881 GATCATCTCATATCAACGCGAGAGACTGACGCGGCGTTTACGCGGGGGTCAACCCCA 6940
Qy 6601 TCCGAGGCAAGCTCTCGCGGAGCGGCACTATGAGGCACTGCTGTGAGACCACTGCAAC 6660
Dh 6941 TCCGAGGCAAGCTCTCGCGGAGCGGCACTATGAGGCACTGCTGTGAGACCACTGCAAC 7000
Qy 6661 ACCCAAGCAAGCGCTATGATGACATGATGATGATGATGATGATGATGATGATGATGAT 6720
Dh 7001 ACCCAAGCAAGCGCTATGATGACATGATGATGATGATGATGATGATGATGATGATGAT 7060
Qy 6721 GTGACTCGGATAGAGTCTGGGCTCAAGTGTGCTTCTGACTCTGCAACCAATGCTC 6780
Dh 7061 GTGACTCGGATAGAGTCTGGGCTCAAGTGTGCTTCTGACTCTGCAACCAATGCTC 7120
Qy 6781 GAAAGAAAGAGGCACTTGAAGCTTGAATACATCAAGAAATATAGTCTCCCAAGAGAG 6840
Dh 7121 GAAAGAAAGAGGCACTTGAAGCTTGAATACATCAAGAAATATAGTCTCCCAAGAGAG 7180
Qy 6841 TTCCCAACAGCTTTACCGGCTGGGCAAGGCTGATTAACCAACCGCTTGTGAAATG 6900
Dh 7181 TTCCCAACAGCTTTACCGGCTGGGCAAGGCTGATTAACCAACCGCTTGTGAAATG 7240
Qy 6901 TGGAAAAGGCGAGATTAACCAACCGGCACTGTTGGGCGGCTGTGCTCTCTCTAG 6960
Dh 7241 TGGAAAAGGCGAGATTAACCAACCGGCACTGTTGGGCGGCTGTGCTCTCTCTAG 7300
Qy 6961 AAAAAGGCGAGGCTCTCCCAAGAGAGGCGCGGCACTGAGGCTTAAGTATGAGACTCAAT 7020
Dh 7301 AAAAAGGCGAGGCTCTCCCAAGAGAGGCGCGGCACTGAGGCTTAAGTATGAGACTCAAT 7360
Qy 7021 GAGATGCGCTTCAACAGGCTGGGCTTGAATGCTTTGGCGAGGCGGCGGCGGCGGAT 7080
Dh 7361 GAGATGCGCTTCAACAGGCTGGGCTTGAATGCTTTGGCGAGGCGGCGGCGGCGGAT 7420

QY 7081 TCAGGCTTTTCACGCGGGGCGGCGCTGCCGATTCGGGACGTACAGCGGCTCTGATGAG 7140
DB 7421 TCAGGCTTTTCACGCGGGGCGGAGCGAGCCGATTCGGGACGTCCGAGCGCCCGCGATGAG 7480
QY 7141 TTGGCCCTTTTGGAGACAGGTTCCATCTTCCATGCCCCCTCGAGGGGAGACTTGA 7200
DB 7481 TTGGCCCTTTTGGAGACAGGTTCCATCTCTCCATGCCCCCTCGAGGGGAGCTTGA 7540
QY 7201 GATCCAGACCTGGAGCGTGAAGAGTGAAGCCCCAACCCCCCGAGGGGGGGTGGCA 7260
DB 7541 GATCCAGACCTGGAGCGTGAAGAGTGAAGCTTAACTCCCCCGAGGGGGGGTGGTA 7600
QY 7261 GCTCCCGGCTCGAGCTCGGGGCTCTGCTCTACTTGTCCGAGAGAGAGCACTCCGTG 7320
DB 7601 ACCCGGGGCTCAGGCTCGGGGCTTGGTCTACTTGTCTCGAGAGAGAGCACTCCGTG 7660
QY 7321 TGTGCTCCATGTCATCTCTGAGACCGGGGCTCTAACTCTCTTGTAGTCCGAGAG 7380
DB 7661 TGTGCTCCATGTCATCTCTGAGACCGGGGCTCTAACTCTCTTGTAGTCCGAGAGAG 7720
QY 7381 GAGAGTTACGATTAACCCCTTGAAGCACTCCCTGTTGGATATCAACAAGGTGTAC 7440
DB 7721 GAAAGTTGCAATGGCCCCCTTGAAGCACTCCCTGTTGGATATCAACAAGGTGTAC 7780
QY 7441 TGTACCAACAACAAGAGCGCTCACTAAGGAGTAAAGTAACTTTGATGAGATCAA 7500
DB 7781 TGTACCACATCAAGAGCGCTCACTAAGGAGTAAAGTAACTTTGATGAGATCAA 7840
QY 7501 GTGCTGACCTCTAAGCACTCACTGTTAAAGAGCACTAAGCTAGCGGCTCCAGAGTC 7560
DB 7841 GCGCTGACGCTCATATGACTGAGTCTTGAAGAGCACTAAGCTAGCGGCTCCAGAGTC 7900
QY 7561 ACCGCAAGGCTCTCAACATGAGAGAGGCTTGCAGTTAACCCCACTTCTGCAAGA 7620
DB 7901 ACCGCAAGGCTCTCACTTTAGAGAGGCTTGCAGTTAACTCCACCTCTGCAAGA 7960
QY 7621 TCTAATATAGGTTTGGGGCTAAGAGAGTCCGAGCTTGTCCGGGAGGCGGTTAAACAC 7680
DB 7961 TCCAGATAGGTTTGGGGCTAAGAGAGTCCGAGCTTGTCCGGGAGAGCGGTTAAACAC 8020
QY 7681 ATCAAGTCCGTGGAAGAGCTTCTGAGAGACTCAGAAAACCAATTTCCCAACCACTT 7740
DB 8021 ATCAAGTCCGTGGAAGAGCTTCTGAGAGACACAAACCAATTTCCCAACCACTT 8080
QY 7741 ATGCGCAAAAATGAGGTTTCTGTGTGAGACCCCAACGAGGGGCAAGAAAGCACTGCG 7800
DB 8081 ATGCGCAAAAATGAGGTTTCTGTGTGAGACCCCAACGAGGGGTAAGAAAGCACTGCG 8140
QY 7801 CTATGCTTAACTCTGACCTCGGGCTCAGGGGTCTGCGAGAGAGATGGCTTTTATGACTT 7860
DB 8141 CTATGCTTAACTCTGACCTCGGGCTCAGGGGTCTGCGAGAGAGATGGCTTTTATGACTT 8200
QY 7861 ACACAAAAATCTCTCAGCGGTGATGAGGGGCTTCTATGATTCAGTATTCGCCGCT 7920
DB 8201 ACACAAAAATCTCTCAGCGGTGATGAGGGGCTTCTATGATTCAGTATTCGCCGCT 8260
QY 7921 CAGCGGTGAGAGTTTCTTGAAGAGCATGGGCGGAAAAAGAGACCTTATGGGTTTTG 7980
DB 8261 CAGCGGTGAGAGTTTCTTGAAGAGCATGGGCGGAAAAAGAGACCTTATGGGTTTTG 8320
QY 7981 TATGATACCGGATGCTTGACTCAACCGTCACTGAGAGAGACATGAGAGCTGAGAGTCC 8040
DB 8321 TATGATACCGGATGCTTGACTCAACCGTCACTGAGAGAGACATGAGAGCTGAGAGTCC 8380
QY 8041 ATATATCGGGCTGCTCTTGGCCGAGAGAGCCCACTGCACTGCTAACTGAG 8100
DB 8381 ATATATCGGGCTGCTCTTGGCCGAGAGAGCCCACTGCACTGCACTGAGAGTCC 8440
QY 8101 AGACTTTAAGTGGAGGCGCTATGTTCAACAGAGAGGCAAACTTGGGTTACAGGCGT 8160
DB 8441 AGACTTTAAGTGGAGGCGCTATGTTCAACAGAGAGGCGCAAGCTTGGGTTACAGGCGT 8500
QY 8161 TGCCGCGCAAGCGGGGTGCTCACCACTAGATGAGGGAACCATCATGCTAGTGA 8220

DB 9501 TGCCGCGCAAGCGGGGTGCTTACCACTAGATGGGGAACACATCATGCTATGTGA 8560
QY 8221 GCTTACGCGGCTTTGAAGTCTGAGGGATATATGCGCCCAATATGCTGTATGCGGAT 8280
DB 8561 GCGTTAGCGGCTTGAAGTCTGAGGGATATATGCGCCCAATATGCTGTATGCGGAT 8620
QY 8281 GACTTGTGCTCATCTCAGAAAGCCAGGGAGACGAGAGAGAGAGAGAGAGAGAGAG 8340
DB 8621 GACTTGTGCTCATCTCAGAAAGCCAGGGAGACGAGAGAGAGAGAGAGAGAGAGAGAG 8680
QY 8341 TTCAAGAGGCTATGACCAAGTATTTGCCCCCTCTGAGTACCCGCCAGAGCGGAGTAT 8400
DB 8681 TTCAAGAGGCTATGACCAAGTATTTGCCCCCTCTGAGTACCCGCCAGAGCGGATAT 8740
QY 8401 GATCTGAGGCTGATTAATCTTGTCTCTCAATGTTGTTGTGGGCTGAGGCGCAAGGC 8460
DB 8741 GACTTGTGCTCATCTCAGAAAGCCAGGGAGACGAGAGAGAGAGAGAGAGAGAGAGAG 8800
QY 8461 GCGCGAGATTAACCTGACCAAGAGACCTTACCACTCAATGCGCGGGGCTGCTGGGA 8520
DB 8801 GCGCGAGATTAACCTGACCAAGAGACCTTACCACTCAATGCGCGGGGCTGCTGGGA 8860
QY 8521 ACAGTTAGACATCCCTGTCATATTCATGCTGGGAAACATATCCAGTACGCGCCGAGC 8580
DB 8861 ACAGTTAGACATCCCTGTCATATTCATGCTGGGAAACATATCCAGTACGCGCCGAGC 8920
QY 8581 ATATGGCTGCAATGCTCTGATGACACATCTTCTTCCATTCATGCTCAAGACAG 8640
DB 8921 ATATGGCTGCAATGCTCTGATGACACATCTTCTTCCATTCATGCTCAAGACAGT 8980
QY 8641 CTGAGACAGAACTCAACTTGAAGTATGAGAGCGGTGATCTCCGAGTCCCTTGAGC 8700
DB 8981 CTGAGACAGAACTCAACTTGAAGTATGAGAGCGGTGATCTCCGAGTCCCTTGAGC 9040
QY 8701 CTCCAGCTTAATTTGAAGGTTAATGAGGCTTGAAGCTTTTCTCTGACACATACAT 8760
DB 9041 CTCCAGCTTAATTTGAAGGTTAATGAGGCTTGAAGCTTTTCTCTGACACATACAT 9100
QY 8761 CCCCAAGACTGACAGGGGTGCTTCAAGCCCTCAAGAAACCTTGGGCGCAACCTCTGAG 8820
DB 9101 CCCCAAGACTGACAGGGGTGCTTCAAGCCCTCAAGAAACCTTGGGCGCAACCTCTGAG 9160
QY 8821 GCGTGAAGAGCCGGGACGTGACGTACAGGCGCTTCAATCTCCGTGGGGGAGAGCG 8880
DB 9161 GCGTGAAGAGCCGGGACGTGACGTACAGGCGCTTCAATCTCCGTGGGGGAGAGCG 9220
QY 8881 GCGGTTGCGGTGATATCTCTTCAATTTGGGCGGTGAAGACCAAGCTCAATCACTCA 8940
DB 9221 GCGGTTGCGGTGATATCTCTTCAATTTGGGCGGTGAAGACCAAGCTCAATCACTCA 9280
QY 8941 TTGCCGAGAGCGGCTCTGATTAATTCAGCTGTTCACGTCGAGCGCGGCGGGGCG 9000
DB 9281 TTGCCGAGAGCGGCTCTGATTAATTCAGCTGTTCACGTCGAGCGCGGCGGGGCG 9340
QY 9001 GACATTTATACAGCGTGTGCTGCGCCGACCCCGCTTATGCTTTGGCTATCTCTA 9060
DB 9341 GACATTTATACAGCGTGTGCTGCGCCGACCCCGCTTATGCTTTGGCTATCTCTA 9400
QY 9061 CTTTGTGAGGGGTAGGCTTTTCTACTACCTCCGCTCG 9099
DB 9401 CTTTGTGAGGGGTAGGCTTTTCTACTACCTCCGCTCG 9439

RESULT 5
ABK8904
ID ABK8904 standard; cDNA; 9678 BP.
XX
AC ABK8904;
XX
DT 07-OCT-2002 (first entry)
XX
DE Human HCV-related polypeptide encoding cDNA.

XX Hepatitis C virus; human; gene; ss; virucide; gene therapy; HCV;
KW fulminant hepatitis C.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT CDS 341..9442
FT /tag= a
FT /product= "Human HCV-related protein"
XX
XX JP2002171978-A.
XX
XX 18-JUN-2002.
XX
XX 01-DEC-2000; 2000JP-00367365.
XX
XX 01-DEC-2000; 2000JP-00367365.
XX
XX (TOKR-) ZH TOKYOITO RINSHO IGAKU SOGO KENKYUSHO.
XX (TORA) TORAY IND INC.
XX
XX MPI: 2002-569884/61.
XX P-PSDB: ABG30688.
XX
XX A gene of a fulminant hepatitis C virus strain and the encoded
PT polypeptide useful in gene therapy to treat hepatitis C.
XX
XX Claim 5; Page 13-25; 36pp; Japanese.
XX
XX The invention relates to a human polypeptide related to hepatitis C virus
CC (HCV), and the polynucleotide encoding it. The polypeptide can be used
CC for the development of gene therapy on fulminant hepatitis C. This
CC sequence represents cDNA encoding a human HCV-related polypeptide
XX
SQ Sequence 9678 BP; 1948 A; 2915 C; 2723 G; 2092 T; 0 U; 0 Other;
Query Match 82.6%; Score 7513.4; DB 6; Length 9678;
Best Local Similarity 89.1%; Pred. No. 0;
Matches 8108; Conservative 0; Mismatches 991; Indels 0; Gaps 0;
QY 1 ATGAGACCAATTCCTTAACCTTAAGAAAAACCAAAACCAACCAACCGTCCGCCACA 60
DB 341 ATGAGACCAATTCCTTAACCTTAAGAAAAACCAAAACCAACCAACCGTCCGCCACA 400
QY 61 GACGTTAAGTTCCCGGCGCGCGCGAGATCGTTGCGGAGTATCTTGTGCGCGCAGG 120
DB 401 GACGTTAAGTTCCCGGCGCGCGCGAGATCGTTGCGGAGTATCTTGTGCGCGCAGG 460
QY 121 GCGCCCAAGTTGGGTGTGCGCGCGCAAGAGAAAGACTTTCGAGCGGTCCAGCCACTGGA 180
DB 461 GCGCCCAAGTTGGGTGTGCGCGCGCAAGAGAAAGACTTTCGAGCGGTCCAGCCACTGGA 520
QY 461 GCGCCCAAGTTGGGTGTGCGCGCGCAAGAGAAAGACTTTCGAGCGGTCCAGCCACTGGA 520
DB 181 AGCGCCCAAGTTGGGTGTGCGCGCGCAAGAGAAAGACTTTCGAGCGGTCCAGCCACTGGA 240
QY 181 AGCGCCCAAGTTGGGTGTGCGCGCGCAAGAGAAAGACTTTCGAGCGGTCCAGCCACTGGA 240
DB 521 AGCGCCCAAGTTGGGTGTGCGCGCGCAAGAGAAAGACTTTCGAGCGGTCCAGCCACTGGA 580
QY 241 TACCCCTGGCCCTTATACGGAATAGGGAATCGGCTGGGAGAGATGCTCTGCTCCGCC 300
DB 581 GCGCCCTGGCCCTTATACGGAATAGGGAATCGGCTGGGAGAGATGCTCTGCTCCGCC 640
QY 301 CGAGGTTCCCGTCCCTCTTGGGGCCCAATGACCCCGGCAATAGTGGCGCAACGTGGGT 360
DB 641 CGAGGTTCCCGTCCCTCTTGGGGCCCAATGACCCCGGCAATAGTGGCGCAACGTGGGT 700
QY 361 AAGGTCATGATACCTTAACGTCGAGTGGCTTGGCGACTCATGGGGTATATCCCTGCTGTG 420
DB 701 AAGGTCATGATACCTTAACGTCGAGTGGCTTGGCGACTCATGGGGTATATCCCTGCTGTG 760
QY 421 GCGCGCCCGCTCGGCGCGCTCGCAGAGCTCTCGCGCATGGGCTGAGAGTCTTGGAGGAC 480
DB 761 GCGCGCCCGCTCGGCGCGCTCGCAGAGCTCTCGCGCATGGGCTGAGAGTCTTGGAGGAC 820

QY 481 GGGGTTAATTTTGCACAGGAACTTACCCTGGTGTCTCTTTCTATCTTTCTTGCTGACC 540
DB 821 GGGGTTAATTTTGCACAGGAACTTACCCTGGTGTCTCTTTCTATCTTTCTTGCTGACC 880
QY 541 CTGCTGTCTGATGACCAACCCCGGTCTCCGTCGCGGAGTAAAGAACTCATGTAACCGGC 600
DB 881 CTGCTGTCTGATGACCAACCCCGGTCTCCGTCGCGGAGTAAAGAACTCATGTAACCGGC 940
QY 601 TACATGATGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 941 TACATGATGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1000
QY 661 CTCACAGTCCCGGCGGTGCTCCGTCGCGGAGAAAGTGGGGAATGATCTCATGTCGTGATA 720
DB 1001 CTCACAGTCCCGGCGGTGCTCCGTCGCGGAGAAAGTGGGGAATGATCTCATGTCGTGATA 1060
QY 721 CCGGTCTCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 1061 CCGGTCTCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1120
QY 781 CACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 1121 CACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1180
QY 841 TCGGTGGGGTGAATGCTCGAGCCCAATGTTTCTTCTGCTCGCGGAGACCACTGCTTT 900
DB 1181 TCGGTGGGGTGAATGCTCGAGCCCAATGTTTCTTCTGCTCGCGGAGACCACTGCTTT 1240
QY 901 GTCCAGAT 960
DB 1241 GTCCAGAT 1300
QY 961 GACATGAT 1020
DB 1301 GACATGAT 1360
QY 1021 CCGGAGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1361 CCGGAGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1420
QY 1081 TACTTCTATGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
DB 1421 TACTTCTATGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1480
QY 1141 GTGAGCGCGGCAACCTATCTGTTGGGGTTCGCGCGGAGCAACCGAGGCGCTCACCC 1200
DB 1481 GTGAGCGCGGCAACCTATCTGTTGGGGTTCGCGCGGAGCAACCGAGGCGCTCACCC 1540
QY 1201 AGCTTATTTGATGATGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
DB 1541 AGCTTATTTGATGATGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1600
QY 1261 CACATCAACCGGCAACCGGCTTCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1601 CACATCAACCGGCAACCGGCTTCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1660
QY 1321 CTGTTTCAACCCCAAGCTTCACTGTCAGATGTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
DB 1661 CTGTTTCAACCCCAAGCTTCACTGTCAGATGTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1720
QY 1381 AGTATGAGGAGCTTCCGGGTGGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
DB 1721 AGTATGAGGAGCTTCCGGGTGGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1780
QY 1441 CCAAGGATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
DB 1781 CCAAGGATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1840
QY 1501 GCGAGAGCTGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
DB 1841 GCGAGAGCTGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1900
QY 1561 ACCGACAGGCTTGGAGCGCCCACTTACAGTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620

Db 1901 ACCGACGAGCTGGAGTGGCCACCTACACATGGGGAGAGATGACAGATGTCTTCTA 1960
Qy 1621 TTGAACAGCACTGACACACCGCTGGGGTCAATGTTGGGCTGACAGTGGAGTGAATCTTCT 1680
Db 1961 CTGAACAGCACCCGACCGCGCAGGGCTCAATGTTGGGCTGACAGTGGAGTGAATCTTCACT 2020
Qy 1681 GGCTACACAAAGACTTGGGGGACACACCGCTGGGCTAGAGCTGACTTCAACGGCAGC 1740
Db 2021 GGTTCACCAAGACTTGGGGGCGCCACCTTGGCGCACAGAGCTGACTTCAACGCCACG 2080
Qy 1741 ACCGACCTGTTGGCCCCACGACCTGTTTGAAGAGCATCTGATACACTTACCTTCAAA 1800
Db 2081 ACCGACTGTGTGGCTTACGGAATGTTTGAAGAGCATCTGATGCACTTATATTAAG 2140
Qy 1801 TGGGGCTTGGGGCCCTGGCTACACGCAAGTGGCTGATCGACTACCTTACAGGCTTGG 1860
Db 2141 TGTGGTCTGGGGCCCTGGCTACACCAAGTGGCTGGTCCACTACCTTACAGACTTGG 2200
Qy 1861 CATTAACCCCTGCAAGTTAATATACATCTTCAAAATTAAGAGTATGTGGAGGGGGTT 1920
Db 2201 CATTAACCCCTGCAAGTTAATATACATCTTCAAAATTAAGAGTATGTAGGGGGGGTT 2260
Qy 1921 GAGCACAGGCTCACGGCTGACATGCAATTTGACTGTGGGATCGTTGCAACTTGGAGAGC 1980
Db 2261 GAGCACAGGCTCACGGCGCGATGCACTTCACTGTGGGATGCGCTGCACTTGGAGAGC 2320
Qy 1981 AGGACAGAAAGTCAACTGTCTCTTGTGTGCACTCCACACAGAAATGGGCAATTTAAGCT 2040
Db 2321 AGGACAGAAAGTCAAGCTGTCTCTCTGTGCACTCTTCAACAGAAATGGGCAATCTCC 2380
Qy 2041 TGGCTTACTCGGACCTGGCCGCTGTGCACTGGTGTCTTCCACCTTCAACCAAAATCT 2100
Db 2381 TGAACCTACTCAACTTACCCGCTTGTCACTGGTGTCTTCTCACTTCAACAGAAATCT 2440
Qy 2101 GTGACGCTACAAATTCATGATAGGCTTACCTGCTTCACTGCTTCAAAATATCATGTCCAGT 2160
Db 2441 GTGACGCTACAAATTCATGATAGGCTTCACTGCTTCACTGCTTCAAAATATCATGTCCAGT 2500
Qy 2161 GAGTGGTAAATCTCTTATTCCTGCTTGAAGGAGCGCCAGGGTTCGGCTGTTATGG 2220
Db 2501 GAGTGGTGTACTTATTCCTGCTTGAAGGAGCGCCAGAGTCTGCGCTGTTATGG 2560
Qy 2221 ATGCTATCTGTTGGGCGAGCGGAGGACCTAGAGAGCTGCTTGGAGAGCT 2280
Db 2561 ATGCTATCTGTTGGGCGAGCGGAGGACCTAGAGAGTGTGCTTGGAGAGCT 2620
Qy 2281 GCGAGCGAGCTAGCTGCAATGCTTCTTATTTTGTCACTTCTTGTGCTGCTTGG 2340
Db 2621 GCGAGTGGCTAGCTGCAATGCTTCTTATTTTGTCACTTCTTGTGCTGCTTGG 2680
Qy 2341 TACATCAAGGTTGGGATGTGCTTCTTGAAGTAACTTCTTCACTGCTGCTTGGCTT 2400
Db 2681 CACATCAAGGTTGGGATGTGCTTCTTGAAGTAACTTCTTCACTGCTGCTTGGCTT 2740
Qy 2401 AGGCTACTGCTTGAAGTGGCCCAAGGCTTATGCTTATAGAGATCTGAGATGCGC 2460
Db 2741 TGGCTACTGCTTATGAGACTGCCCCGAGGCTTATGCTTATAGAGATCTGAGATGCGC 2800
Qy 2461 CAGATAGAGCGGCTCTGCTGATATGATCACTTCTTATCTCTCACTCCCGGGATATAG 2520
Db 2801 CAGATAGAGCGGCTTGTGATATGATCACTCCCTTCACTCACTCACCCGGGATATAG 2860
Qy 2521 ACCCTTCTCAAGCGGTTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
Db 2861 ACCCTTCTCAAGT 2920
Qy 2581 GTCCAGAGTGGGACCACTATGACAGTGGCGGCTGCTGATGAGTATGAGTATGAGTATGAGT 2640
Db 2921 ATTCAAGAGTGGGATCACTCACTGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2980
Qy 2641 GTGCAATATTTCAACCAAGT 2700

Db 2981 GTCACTATATCTGCCCCGGGT 3040
Qy 2701 GGGCTGCTTACCTCTTAAAGGTGTTTGAACGGGCTGCGGCTTCTTGTGACGGGCTCAC 2760
Db 3041 GGGCTGCTTACCTCTTAAAGGTGTTTGAACAGTGTGCGGCTTGTGTGTGTGTGTGTGTGTGTGT 3100
Qy 2761 GCTCTACTGAGATGTGCAACATGACAAAGGATCTCCGGGGGGGAGGATAGTCAAGT 2820
Db 3101 GCTCTGATTAAGGATGTGCGCTTGT 3160
Qy 2821 GCGCTACTAGCCCTTGGAGAGTGTGATGCACTTATCATATGACACTTACCTCATG 2880
Db 3161 GCGCTATTTGGCCCTTGGAGAGTGTGACCTGACCTATCATATGACACTTACCTCATG 3220
Qy 2881 TCGGATTTGGGCTGTAGTGGCTTGGGGAGCTGGGGGTCCCGGTTGAGCCCTATATCTTC 2940
Db 3221 TCGGACTGGGCGCTGAGCGGCTGCGGACTTGTAGCGGTCCCGGGAACCATATCTTC 3280
Qy 2941 AGTCCGATGGAAGAAAGTCAATTTGTCTGGGAGAGGAGACAGCTGCTTGGGGACATT 3000
Db 3281 AGTCCGATGGAAGAAAGTCAATTTGTCTGGGAGAGGAGACAGCTGATGTGGGGACATT 3340
Qy 3001 TTAACGGAATCTCCGTGTCGCGCCGACTTGTGTGGAGGTCTCTTGGCCGACTGAT 3060
Db 3341 CTACATGACTTCCGCTGTGCGCCGACTGCGCCAGAGATCTCTCTGCGCCGACTGAT 3400
Qy 3061 GGGCTATCTTCAAGGGGTGTGATCTTGTGCCCCCATGACTGTCTTACGCCCCAGAGACA 3120
Db 3401 GGGCTACACTTCAAGGGGTGTGATCTTGTGCCCCCATGACTGTCTTATGCCCCAGAGACA 3460
Qy 3121 CGTGGCTTTTGGGACCAATAGTGTGATGATGACATGCGGGGCGGCAAGAGAGAGCT 3180
Db 3461 CGAGGCTCTTGGGCGGCAATGT 3520
Qy 3181 GGGGAAATTCAGGTCTGTGTCACAGTCACTGACTCTTCTTGGAAATCATCTGTGGG 3240
Db 3521 GGGGAAATTCAGAAATCTGTGTCAAGTCTTCAAGTCTTCTTGGAAATCATCTGTGGG 3580
Qy 3241 GTTTTGTGATCTGT 3300
Db 3581 GTTTTGTGATCTGT 3640
Qy 3301 GTCACGAGATGTATCTCAAGTGTGAGGGGAGCTTATAGTGGTGGCCAGCCCTGAGG 3360
Db 3641 GTCACGAGATGTATCTCAAGTGTGAGGGGAGCTTATAGTGGTGGCCAGCCCTGAGG 3700
Qy 3361 ACTTAAATCTTGTGAGCGGTGTGACAGTGTGAGCGGTGTGACTGTATCTGTACGCGGAC 3420
Db 3701 ACCAAGTCTTGTGAGCGGTGTGAGCGGTGTGAGCGGTGTGACTGTATCTGTACGCGGAC 3760
Qy 3421 GCTGATGTCTATCCCGCTGTGAAAGCGGGGAGCAAAACGGGAGCGCTACTCTCCGAGA 3480
Db 3761 GCTGATGTCTATCCCGCTGTGAAAGCGGGGAGCAAAACGGGAGCAATGCTCTCCGAGA 3820
Qy 3481 CCTCTTTCACCTTGAAGGGGTCTTCAAGAGGCGCGGTGTGTATGCCCCAGGGGCAAGCT 3540
Db 3821 CCTCTTTCACCTTGAAGGGGTCTTCAAGAGGCTTCAAGAGGCTTCAAGAGGCTTCAAGAGG 3880
Qy 3541 GTGGAATCTTCCGGGCAAGT 3600
Db 3881 GTTGGGCTTCTTCCGAGCAGGT 3940
Qy 3601 CCGGTTGAGACATTCGACATGCGAGCGGTCCCGCACTTATGTGACAAACAGACACCA 3660
Db 3941 CCGGTTGAGACATTCGACATGCGAGCGGTCCCGCACTTATGTGACAAACAGACACCA 4000
Qy 3661 CCTGCTGTGCCCCAGACTATCAGTGTGAGTCTTGTGATGCCCCGAGCTGCAAGTGAAG 3720
Db 4001 CCGGCTGTGCCCCAGACTATCAGTGTGAGTCTTGTGATGCCCCGAGCTGCAAGTGAAG 4060
Qy 3721 AGCACCAAAAGTCTGT 3780
Db 4061 AGCACCAAAAGTCTGT 4120

QY	3781	TCAGTGGCTGCACCCCTGGGGTTTGGGGCGTACTGTGTAAAGGACATGGCATTAATCCC	3840
Dp	4121	TCGGTAGTGCACCCCTGGGGTTTGGGGCGTACTGTGTAAAGGACATGGCATTAATCCC	4180
QY	3841	AACATTAGAGTCTGAGAGTCAGGACTGTGACGACCGGGGCGCCCATCAGTACTCCACATAT	3900
Dp	4181	AACATTAGAGTCTGAGAGTCAGGACTGTGACGCGGGAGGCGCATCAGTACTCCACATAT	4240
QY	3901	GGCAAAATTCCTGCGCGATGGGGGCTGTGCGGGCGGCGCTACGACATCATCATGTGAT	3960
Dp	4241	GGCAAAATTCCTGCGCGATGGGGGCTGTGCGGGCGGCGCTATGACATCATCATATGAGAT	4300
QY	3961	GAATGGCATGCGTGGGACTTACCAACCATCTTGGGATCGGGAACAGTCTTTGATCAAGCA	4020
Dp	4301	GAATGGCATGCGTGGGACTTACCAACCATCTTGGGATCGGGAACAGTCTTTGATCAAGCA	4360
QY	4021	GAGACAGCTGGGGTCAAGCTAACTGTGTGGCTACAGTACGCCCCCTGGTCAGTACA	4080
Dp	4361	GAGACAGCTGGGGTCAAGCTAACTGTGTGGCTACAGGACACCCCCCGGTCAGTACA	4420
QY	4081	ACCCCCCAACCCCACTATGAGAGAGTGGCCCTTGGGCGAGGAGGGCGAGATCCCTTCTAT	4140
Dp	4421	ACCCCCCATCCCGATATAGAAAGAGTGGCCCTCGGGCGGGAGGTGAGATCCCTTCTAT	4480
QY	4141	GGGAGGGGGGATTTCCCTGTCTTTATCATCAAGGAGGAAGAGCATCTGATCTTCTGCATTCA	4200
Dp	4481	GGGAGGGGGGATTTCCCTATCTCGATCAAGGAGGAAGAGCATCTGATTTCTGCACATCA	4540
QY	4201	AAGAAAAAGTGTGACGAGCTCGCGCGGCGCTTGGGGTATGGGCTTGAACTCAGTGGCA	4260
Dp	4541	AAGAAAAAGTGTGACGAGCTCGCGCGGCGCTTGGGGGCAATGGGCTTGAAATGCGTGGCA	4600
QY	4261	TACTACAGAGGGTTGAGCGTCTCCGTATATCCAACTCAAGGAGACGTATGTGTGTGCC	4320
Dp	4601	TACTATAGAGGGTTGAGCGTCTCCATATATACAGCTCAGGAGAGATGTGTGTGTGTGCC	4660
QY	4321	ACCGAGCGCCCTCATAGCAGAGGGTATATCTGGGGGACTTTGATCCGGATGTGACTGCAAGTA	4380
Dp	4661	ACCGAGCGCCCTCATAGCAGGGGTACCTGGAAGACTTTGATCTCGATGTGACTGCAAGTA	4720
QY	4381	GCGGTCACTCAAGTTGTAGACTTCAGTTTGTAGACCCCATTCACCATTAACCAACAGATT	4440
Dp	4721	GCGGTCACTCAAGCTGTGACTTCAGCTTGGACCCCATCTTCATATTAACCAACAGACT	4780
QY	4441	GTCCTCAAGAGCGCTGTCTACAGTACGACGCGCGGGGTGGCAGGGGTATGGGGAAAGACTG	4500
Dp	4781	GTCCTCAAGAGCGCTGTCTACAGGAGTACGCGCGCGGGGCCCAACAGGTAGAGGAAGACAG	4840
QY	4501	GGCATTTATAGGTATGTTTCCACTGCTGTGTAGAGGAGGCTCAGGAATGTTTGAACAGTGTAGT	4560
Dp	4841	GGCATTTATAGGTATGTTTCCACTGTGTAAAGAGCTCAGGAATGTTTGAACAGTGTAGT	4900
QY	4561	CTCTGTGAGTGTACGAGCGCAGGGGCGCATGTGTATGAGCTCACACCATCGGAGACCAACC	4620
Dp	4901	CTTTGTGAGTGTATGAGCGAGGGGCTGTGTGTATGATCTTCACACGCGGAGACCAACC	4960
QY	4621	GTCAGGCTCAAGGCGGTATTTCAACGCGCCGGTTTGCTGTGTGCGAAAGACATCTTGTAG	4680
Dp	4961	GTCAGGCTTAGAGCTATTTCAACACGCGCGGCTTACCCGGTGTCAAGAACATCTTGTAA	5020
QY	4681	TTTTGGGAGGCGAGTTTTCACCGGGCTCAACACATATGACGCCCATCTTCTCCCAACA	4740
Dp	5021	TTTTGGGAGGCGAGTTTTCACCGGGCTCAACACATATGACGCCCATCTTCTCCCAACA	5080
QY	4741	AAGCATTCGGGGGAAAAATTTGCGATATCTTAAACAGCTACAGGCTACAGTGTGCGTAGG	4800
Dp	5081	AAGCAAGCGGGGGAACATTGGCGGTACCTAAGTAGCTCAAGAGTACGCTGTGGCGCAGA	5140
QY	4801	GCCAAAGCCCCCCCCGCTCTGTGGGAGCTCATGTGAAAGTGTTTGATCTGCATCAAGCCC	4860
Dp	5141	GCCAAAGCCCCCCCCGCTCTGTGGGAGCGCATGTGAAAGTGTCTGTGGCCGACTCAAGCTC	5200

OY	4861	ACA CTGGTGGAGCCCAACCTCTCTGTACCGGTGGGGCTCTGTATACCAACGAGGTCAAC	4920
Db	5201	ACGCTTCGGAGCCCAACCTCTCTCTGTACCGGTGGGGCTCTGTATACCAACGAGGTCAAC	5260
OY	4921	CTCACATCATCCCGTGAAGAAATACATGCAACCTGCATGCAAGCCGACTTGAAGTCAATG	4980
Db	5261	CTCACATCATCCCGTGAAGAAATACATGCAACCTGCATGCAAGCCGACTTGAAGTCAATG	5320
OY	4981	ACCAGCAATGGGTCTTTGGCAGGGGAGTCTTGGCGGCGCTCGCCGGTATTTGCGTGGCG	5040
Db	5321	ACCAGCAATGGGTCTTTGGCAGGGGAGTCTTGGCGGCGCTCGCCGGTATTTGCGTGGCG	5380
OY	5041	ACCGGGGTGTTTGCATCATCGGCGGCTTGCACTTAACAGCAGGCGGTCTTGTCGCGC	5100
Db	5381	ACTGGAATGCTTTTCATCATCGGCGGCTTGCACTTAACAGCAGGCGGTCTTGTCGCGC	5440
OY	5101	GACAAAGAGTCTCTATGAGCTTTTGTATGAGATGAGAGAAATGTGCTCTAAGCGGCT	5160
Db	5441	GATAAAGAGTCTCTATGAGCTTTTGTATGAGATGAGAGAAATGTGCTCTAAGCGGCT	5500
OY	5161	CTCATTTGAAGAGGGGACGCGGATAGCCGAGATGCTGAATTCGAATCCAAAGCTTATTG	5220
Db	5501	CTCATTTGAAGAGGGGACGCGGATAGCCGAGATGCTGAATTCGAATCCAAAGCTTCTG	5560
OY	5221	CAGGAACCTTCCAAACAAGCTCAAGACATCAACCCACTGAGAGGCTTATAGGCCGACG	5280
Db	5561	CAGGAAGCTCTTAAAGCAGGCCCAAGAAATCAACCCGCTATGAGGCTTATAGGCCCA	5620
OY	5281	GTAGAACAATTTCTGGGCAACAACATGTGAACTTTAGCGGACTCCAAATCCTCGCA	5340
Db	5621	GTAGAACAATTTCTGGGCAACAACATGTGAACTTTAGCGGACTCCAAATCCTCGCA	5680
OY	5341	GGACTTATCAACACTGCCAGAGAACTCTGCAGTAGCTTTCATGATGCGCTTCAATGCCGC	5400
Db	5681	GGATTTGCAACACTGCGCAGGAAACCCGCGGTGCTTCCATGATGCGATTCAGTGCCGC	5740
OY	5401	CTCACACGATCCGCTGTCAACAAGACCACTATCTTTCTCAACATTTTGGGGGCGCTGCTA	5460
Db	5741	CTCACACGATCCGTTGTGACCAAGTACCACTCTTCTCAACATCATGGGAGGCTGTGTTA	5800
OY	5461	GCATTCCCAATTTGACCAACCGCGGGGGGCACTGGCTTCTGTTCAGTGGCTTATGGGA	5520
Db	5801	GCATTCCCAAGTGCACACACCGCGGGGGCAACCGGCTTTGTCTGTAATGGCTGTGGGG	5860
OY	5521	GCTCCCGTAGGCAGTATAGGCTTAGGTAAAGTGTCTAATGCAATCTTGGCAGGGTATGTT	5580
Db	5861	GCTCCCGTAGGCAGCATAGGCTGGGTAAAGTGTCTGATGCAATCTTGGCAGGATATGTT	5920
OY	5581	GCGGGCATTTGGGGGGCTCTGTGTGCATTCAGATCATATGTCTGGCAGAAAGCCCTCATG	5640
Db	5921	GCGGGCATTTTGGGGGGCCCTGTGTGCATTCAGATCATATGTCTGGCAGAAAGCCCTCATG	5980
OY	5641	GAGGATGTCGCAACTTGCTGCTGCGAATTCGTCTCGGGTGGCTTGGTATGGAGTCT	5700
Db	5981	GAGGATGTCATCATCTACTGCTCTGGGATCTGTCTCCGGAGGCCCTGTGTGTGGGGGTCT	6040
OY	5701	ATCTGCGGCGCATTTCTGCGCGCAACAGTGGGACCGGGGGAAGCGCGCTTCCATGATG	5760
Db	6041	ATCTGCGGCGCATTTCTGCGCGCGCAAGTGGGACCGGGGGAAGCGCGGCTCAATGATG	6100
OY	5761	AATGAGCTCATTTGCTTGTCTTCCAGAGAAATCAAGTCCGCCCAACCTCACTACGTAGC	5820
Db	6101	AACAGGCTTATTTGCTTGTCTTCCAGAGAAACACAGTCCGCCCTTACTCATCTAGTAGC	6160
OY	5821	GAGTCGGAATGGTGGCAGCGGTGACCACTACTCTGAGTCTTCTTACTATACAGCTTA	5880
Db	6161	GAGTCGGAATGGTGGCAGCGGTGACCACTACTCTGAGTCTTCTTACTATACAGCTTA	6220
OY	5881	CTCAGAAAGCTCCAACTGGATTACTGAGACTGCCCATCCGATCGAGCGGCTCGTG	5940
Db	6221	CTCAGAAAGCTCCAACTGGATTACTGAGACTGCCCATCCGATCGGATTCGTGG	6280
OY	5941	CTCCGCGATGTGTGGGACTGGGTTTGACCATCTCTTAACAGCTTTAAATTTGGCTGACC	6000

Db 6281 CTCGCGACGCTGTGGAGCTGGGTTTGCAACATCTTGCAGAGCTTCAAAAATTTGGCTGACC 6340
Qy 6001 TCCAAATTAATTTCCCAAAAGATGCCCGGCTCCCTTTGTCTCTCTCAAAAAGGGTAAAG 6060
Db 6341 TCTAAATTTGTTCCCAAGCTGCCCGGCTCCCTTCACTCTCTCTCAAAAAGGGTAAAG 6400
Qy 6061 GGGGTGTGGGCGGCACTGGCATCATGACCAACGGGTCTTGGCGGCGCAATATCTCT 6120
Db 6401 GGTGTGTGGGCGGCACTGGCATCATGACCAACGGGTCTTGGCGGCGCAACATCTCT 6460
Qy 6121 GGCATGTCCGCTTGGGCTCATGAGATCAACGGGCTTAAGACTTGCATGAATATCTGG 6180
Db 6461 GGCATGTCCGCTTGGGCTCATGAGATCAACAGGGCTTAAGACTTGCATGAACACTGG 6520
Qy 6181 CAGGGGACCTTTCTCTATCATATTGTACAAGGAGGCGCAGTGGCGGCAAAACCCGCGCA 6240
Db 6521 CAGGGGACCTTTCTCTATCATATTGTACAAGGAGGCGCAGTGGCGGCAAAACCCGCGCA 6580
Qy 6241 AACTTTAAGGTCCCATCTGAGAGGTGGCGGCTCAGAGTACGGCGAGTACGACAGAC 6300
Db 6581 AACTTAAAGACCCGCATCTGAGAGGTGGCGGCTCAGAGTACGGCGAGTACGACAGAT 6640
Qy 6301 GGGTCATACCACTACATTAACAGGACTCACCACTGATTAAGTAAAGTCCCTGCGCACTA 6360
Db 6641 GGGTCGATCTCTATGTAAACAGGACTCACCACTGATTAAGTAAAGTCCCTGCGCACTA 6700
Qy 6361 CCTCTCCCGAGTTCTTTCTGGGTGAGACGAGATGACATCCATAGTTTGGCCCCA 6420
Db 6701 CCTTCTCCAGAGTTTCTCTGGGTGAGACGAGTGTGAGATCAAGTATTGACCCACA 6760
Qy 6421 CCGAAGCCGTTTTTCCGGAGTGAAGTCTCGTTCGCTTGGGCTTAATTCATTTGTGCTC 6480
Db 6761 CCGAAGCCGTTTTTCCGGAGTGAAGTCTCGTTCGCTTGGGCTTAATTCATTTGTGCTC 6820
Qy 6481 GGGTCCACGCTTCTTTCGCAACCTTGAACCCGACACAGACGTAATGATCATAGCTTACA 6540
Db 6821 GGGTCCACGCTTCTTTCGCAACCTTGAACCCGACACGTAATGATCATAGCTTACA 6880
Qy 6541 GATTCATCTCATATCAACGGCGAGACTGACGCGCGGCTTTAGCGCGGGGTCACCCCA 6600
Db 6881 GATTCGCGCCCACTCAACGGCGAGACTGACGCGCGGCTTTAGCGCGGGGTCACCCCA 6940
Qy 6601 TCCGAGGCAAGCTCTCGGAGGACGCTATTCGGACATCGCTGGGAGGCACTGAC 6660
Db 6941 TCTGAGGCGAGCTCTCGAGTGAAGCACTATCAACGCTGCGGAGCACTTGAC 7000
Qy 6661 ACCCAGGCAAGCTATGATGTGACATGTGATGCTTAACCTGTTCAATGAGGAGCGAT 6720
Db 7001 ACCCAGGCAAGCTATGATGTGACATGTGATGCTTAACCTGTTCAATGAGGAGCGAT 7060
Qy 6721 GTTACTCGGATTAAGTCTGGGTCGAAAGTGTGTTTGGACTCTCTCGACCCATGTCTC 6780
Db 7061 GTTACTCGGATTAAGTCTGGGTCGAAAGTGTGTTTGGACTCTCTCGACCCATGTCTC 7120
Qy 6781 GAAGAAAGAGGAGCACTTGAAGCTTGCATCAATCAAGATTAACGCTCCCAAGAAAGG 6840
Db 7121 GAAGAAAGAGGAGCACTTGAAGCTTGCATCAATCAAGATTAACGCTCCCAAGAAAGG 7180
Qy 6841 TTCCCAACAGCTTTACCGGCTTGGGCAACGCTGATTAACAACCAACGCTTTGTGAATCG 6900
Db 7181 TTTCACACGAGCTTTACCGGCTTGGGCAACGCTGATTAACAACCAACGCTTTGTGAATCG 7240
Qy 6901 TGGAAAAGGCAATTAACCAACCGGCACTGTGGGAGCTGTCTCTCTCTCTCTCTCTCTCT 6960
Db 7241 TGGAAAAGGCAATTAACCAACCGGCAACGCTGTGGGAGCTGTCTCTCTCTCTCTCTCTCTCT 7300
Qy 6961 AAAAAGGCAAGCT 7020
Db 7301 AAGGCT 7360
Qy 7021 GGAGATGCTCTTCAACAGCTGAGCTGATTAAGTCTCTTGGCAAGCTCCCAAGCGGAGAT 7080

Db 7361 TCAGAAAGCTCTCCAGCAACTGCGCATCAAGACCTTTTGGCAGGCCCCCTCGAGCGTAT 7420
Qy 7081 TCAGAGCTTTTCCACGAGGAGGAGGAGCTGCGGATTTCCGGAGTCAAGAGCTCTCTGATGAG 7140
Db 7421 GCAAGCTCTGTCACGAGGAGGAGGAGCTGCGGATTTCCGGAGTCAAGAGCTCTCTGATGAG 7480
Qy 7141 TTGGGCTTTTCCGAGACAGGTTTCAATCTCTTCAATGAGCCCCCTCGAGGAGGAGCTTGA 7200
Db 7481 CCGGCCCCCTCAGAGAGAGGTTTCCGCTCTCTCTATGAGCCCCCTCGAGGAGGAGCTTGA 7540
Qy 7201 GATTCAGACCTGAGGCTGAGAGTGAAGAGCCCCCAACCCCCCCCCCGAGGAGGAGTGA 7260
Db 7541 GATTCGAGACCTGAGGCTGAGTGAAGAGCTTCAACCTCTCCCGCAGGAGGAGGAGTGA 7600
Qy 7261 GCTCCCGGCTGAGGAGCTGAGGAGCTGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 7320
Db 7601 GCTCCCGGCTGAGGAGCTGAGGAGCTGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 7660
Qy 7321 TGCTGCTCAATGATCACTCTGAGACCGGAGCTTAAATACTCTTGTAGTCCGAAAGAG 7380
Db 7661 TGCTGCTCAATGATCACTCTGAGACCGGAGCTTAAATACTCTTGTAGTCCGAAAGAG 7720
Qy 7381 GAGAGTACCGATTAACCCCTTGAAGCACTCCCTGTGGATATCAACAAGGTATC 7440
Db 7721 GAAAAGTTCGAATCAACCTTGAAGTACCTGCTGTGGATACATTAACAAGGTATC 7780
Qy 7441 TGTACCAACAAGAGGCGCTCACTAAGGCTTAAAGGTAAGTATCTTGTAGATGCA 7500
Db 7781 TGTACCAACAAGAGGCGCTCACTAAGGCTTAAAGGTAAGTATCTTGTAGATGCA 7840
Qy 7501 GTGCTGACCTCTACTACGACTCACTTAAAGGATTAAGTACGAGGCTCAAGGTC 7560
Db 7841 GTGCTGACGCTCACTACTATGATCACTCACTTAAAGGATTAAGGATCAAGGTCCTCAAGGTC 7900
Qy 7561 ACCGCAAGGCTCTCAACATGAGAGAGCTTGCAGTTAACCCCACTTCTGACAGA 7620
Db 7901 AGCGCAAGGCTCTCAACATGAGAGAGCTTGCAGTTAACCCCACTTCTGACAGA 7960
Qy 7621 TCTAAATATAGGTTTGGGCTAAGAGTCCGAGCTTGTCCGGAGGAGGCGCTTAAACAC 7680
Db 7961 TCCAAAGTATGATTTGGGCGCAAGAGGCTCGAGCTTGTCCGGAGGAGGCGCTTAAACAC 8020
Qy 7681 ATCAAGTCGTTGGAAGAGCTCTGAGAGACTCAAGAAACAACAATTTCCCAACCAT 7740
Db 8021 ATCAAGTCGTTGGAAGAGCTCTGAGAGACTCAAGAAACAACAATTTCCCAACCAT 8080
Qy 7741 ATGGCAAAAATAGAGGTGTTCTGCTGAGACCCCAACGAGGAGGAGCAAGAGCTGCG 7800
Db 8081 ATGGCAAAAATAGAGGTGTTCTGCTGAGACCCCAACGAGGAGGAGCAAGAGCTGCG 8140
Qy 7801 CTATGCTTTAACCTTGAACCTCGGCGTCAAGGCTTGGAGAGATGAGCCCTTTATGACATT 7860
Db 8141 CTATGCTTTAACCTTGAACCTCGGCGTCAAGGCTTGGAGAGATGAGCCCTTTATGACATT 8200
Qy 7861 ACACAAAACCTTCTCAGGCGGATGAGGAGGCTTCTTATGATTTCCAGTATTCGCCGCT 7920
Db 8201 ACACAAAACCTTCTCAGGCGGATGAGGAGGCTTCTTATGATTTCCAGTATTCGCCGCT 8260
Qy 7921 CAGCGGATGAGGTTTCTTTGAAGAATGGGCGAAGAAAGAGACCTTATGAGGTTTTCG 7980
Db 8261 CAGCGGATGAGGTTTCTTTGAAGAATGGGCGAAGAAAGAGACCTTATGAGGTTTTCG 8320
Qy 7981 TATGATTAACCGAGCTTTGATCTAACGCTACGTAAGAGAGCATCAAGAGCTGAGGATCC 8040
Db 8321 TATGATTAACCGAGCTTTGATCTAACGCTACGTAAGAGAGCATCAAGAGCTGAGGATCC 8380
Qy 8041 ATATATCGGCTGCTCTTGTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8100
Db 8381 ATATATCGGCTGCTCTTGTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8440
Qy 8101 AGACTTTAGTGGAGGAGGCTTATGTTCAACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 8160
Db 8441 AGACTTTAGTGGAGGAGGCTTATGTTCAACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 8500

OY	81.6	TTGCCGCGCAGCGGGGTCCTACCACTAGATGAGGAAACACATTCACATGCTACGTAA	8220
Db	8501	TGCGGCGCAGCGGGGTCTAACCACTAGATGAGTAACACATTCACATGCTAATGTA	8560
OY	8221	GCCTTAGCGGCTGTAAAGCTGCAGGGAATATCGCGCCCAATGCTGTATGCGCGAT	8280
Db	8561	GCCTTAGCGGCTGTAAAGCTGCAGGGAATATGTTGGGCCCAATGCTGTATGCGCGAT	8620
OY	8281	GACTTTGGTGTTCATCTCAGAAAGCCAGGGAGCCGAGAGAGACAGCGGAACTTAAAGCC	8340
Db	8621	GACCTTAGTGTTCATCTCAGAAAGCCAGGGAGCTGAGAGAGACAGCGGAACTTAAAGCC	8680
OY	8341	TTCAACGAGGCTATGACCAAGGTATTTCTGCCCTCTGGTGAACCCGCCAGACCGAGTAT	8400
Db	8661	TTCAACGAGGCGCATGACCAAGGTATCTGCGCCCTCTGGTGAATCCCCCAGACCGGAATAT	8740
OY	8401	GATCTGAGCTGATTAACATCTTGTCTCTCAATGTGTCTGTGGCGCTGGGCCAAGAGC	8460
Db	8741	GACCTGAGCTAATTAACATCTCTGTTCTCAATGTGTCTGTGGCGTTGGCGCGCGGCGC	8800
OY	8461	CGCGCGAGATACCTACCTGACCAAGAGACCTTACCAATCGCCCGGGCTGCTGGGAA	8520
Db	8801	CGCGCGAATATCTACCTGACCAAGAGACCAACCACTCCACCTCGCCGCGCTGCTGGGAA	8860
OY	8521	ACAGTTAGACACTCCCTGTCAATTCAATGCGTGGGAAACATCATCAGTACGCCCGACC	8580
Db	8861	ACAGTTAGACACTCCCTGTCAATTCAATGCTGGGAAACATCATCAGTATGCTTCAAC	8920
OY	8581	ATATGGGCTCGCATGTGCTGTATGACACACATCTTCTTCTCATGTCATNGCTCAAGACAG	8640
Db	8921	ATATGGGTTGGCATGTGCTGTATGACACACATCTTCTTCTCATGTCATNGCTCAAGACAC	8980
OY	8641	CTGGACCAGAACCTCACTTTGAGATGTACGAGCGGCTGTACTCGTGAATGCCCTTGAC	8700
Db	8981	CTGGACCAGAACCTCACTTTGAGATGTATGATCAGTATACCTCGTGAATCCTTTTGAC	9040
OY	8701	CTCCAGCTATAATTGAAGGTTGATNGGCGTGAACGTTTTCTCTGCAACAADACACT	8760
Db	9041	CTTCAGCATATAATTGAAGGTTGACGCGCTTGACGCTTTTCTATGACAACAADACTCT	9100
OY	8761	CCCAACGAATGACACAGGGTGGCTTCAGCCCTCAGAAAACTTGGGGCCACACCCCTCAGA	8820
Db	9101	CACACAGAACTGACCGCGGGTGGCTTCAGCCCTCAGAAAACTTGGGGCGCCACCCCTCAGG	9160
OY	8821	GCGTGAAGAGCCGGGCACTGTACAGTCAAGGCGTCTTCATCTCCGTTGGGGGAGAGCG	8880
Db	9161	GTTGTGAAGAGCTCGGCTCTCGCGCAGTCAAGGCGTCTTCATCTCCGTTGGGGGAGAGCG	9220
OY	8881	GCGGTTGGCGGTCAATCTCTTCAATTGGGGGGTGAAGACCAAGCTCAAACTCACTCA	8940
Db	9221	GCGGTTGGCGGTCAATCTCTTCAATTGGGGGGTGAAGACCAAGCTCAAACTCACTCACTCA	9280
OY	8941	TTGCCGAAGGCGCGCTCTCGATTATTCAGCTGTTCAACGTCGCGCGCGCGGGGCG	9000
Db	9281	TTGCCGAAGGCGCGCTCTCGATTATTCAGCTGTTCAACGTCGCGCGCGCGGGGCG	9340
OY	9001	GACATTTATCACAGCTGTCCGCTGCCAGACCCGCTTATGCTTTGGCTTACTCTTA	9060
Db	9341	GACATTTTTCACAGGCTGTCCGCGCGCCGACCCGCTATTAATCTTTCGGCTACTCTTA	9400
OY	9061	CTTTTGTAGAGGTAGGCTTTTCTTACTCCCGGCTCGG	9099
Db	9401	CTTTTGTAGAGGTAGGCTCTTCTTCTACTCCCGGCTCGG	9439

RESULT 6
AAQ38221
AAQ38221 standard, cDNA to mRNA, 9511 BP
ID
XX
AC
XX
DT
27-AUG-2003 (revised)

Dt	25-MAR-2003	(revised)
Df	01-JUL-1993	(first entry)
XX		
XX	Non A non B hepatitis virus; plasma; degenerate; ss.	
XX		
XX	Non-A.	
OS	non-B hepatitis virus.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	342..9443
FT		/tag= a
XX		
PN	EP532167-A2.	
PD	17-MAR-1993.	
XX		
PF	30-JUL-1992;	92EP-00306952.
XX		
PR	09-AUG-1991;	91JP-00287402.
PR	05-DEC-1991;	91JP-00360441.
XX		
PA	(IMMO) IMMUNO JAPAN INC.	
PI	Okamoto H, Nakamura T;	
DR	WPI, 1993-087166/11.	
DR	N-PsDB; AAR33539, AAR33214.	
XX		
PT	Polynucleotide(s), polypeptide(s) and antibodies of NANBH virus - useful for detecting NANBH, as a vaccine and for screening blood samples.	
PS	Claim 7; Page 59-64; 93pp; English.	
CC	RNA was isolated from the plasma of human patients positive for NANBH virus (strain HC-J8) and was subjected to reverse transcription to produce cDNA. The resulting cDNA was amplified by PCR, and nucleic acid sequences determined by analysis of clones obt'd. by PCR amplification (42 clones in total). The NANBH HC-J8 degenerate genome was found to contain a degenerate open reading frame encoding polypeptide precursors of 3033 amino acid residues. See also AAQ38172-220. (Updated on 25-MAR-2003 to correct PN field.) (updated on 27-AUG-2003 to correct OS field.)	
SQ	Sequence 9511 BP; 2020 A; 2709 C; 2539 G; 2108 T; 0 U; 135 Other:	
<hr/>		
Query Match	62.1%; Score 5652.2; DB 2; Length 9511;	
Best Local Similarity	75.5%; Pred. No. 0;	
Matches 6667; Conservative 104; Mismatches 2128; Indels 0; Gaps 0		
Qy	1 ATGAGCACAATTCCTTAAACCTCAAGAAGAAAACCAAAGAACCAACCGTGCCACAA 60	
Db	342 ATGAGCACAAAACTCTTAACCTCAAGAAAAAACAAGAAACAAACCGGCCACAG 401	
Qy	61 GACGTTAAGTTCCGGGCGCGCGGCAGATCGTTGGCGGAGTAGTACTTGTGTGCGCGACAG 120	
Db	402 GACGTTAAGTTCCGGGCGCGCGGTGATCATCTGTGGCGAGTTACTTGTGCGCGCAAG 461	
Qy	121 GGCCCCAGGTTGGGTGTGCGCGCGCAAGGAAGAACTTGGAGCGGTCCAGACCAGTGG 180	
Db	462 GGCCCCAGGTTGGGTGTGCGCGCGCAAGGAAGAACTTCTGACAGCATCCCAAGCGCGTGA 521	
Qy	181 AGGCGCGAGCCCATCTCTTAAGATGGCGCTCCACTGGAAATTCGGGGGAAAAACAGGA 240	
Db	522 CGAGCGCAGCCATCCCGCAAGAAATGCGCCTCCACCGGAAAGTCTGGGGAAAGCCAGGA 581	
Qy	241 TAACCCCTGGCCCCCTATACCGGATGAGGAGACTCGCTGGGACAGATAGCTCTGTCCCC 300	
Db	582 TATCTTGCGCCCTGTAGAGAAACGAGGGTTGGGCGGGGTTGGCTCTGTCCCCC 641	
Qy	301 CGAGGTTCCCGTCCCTCTTGGGGGCCCAATGACCCCGGAGATAGCTGCGGCAACGTGGGT 360	
Db	642 CGCGGATCTCGCTTAATTGGGGGCCCAACGAGCCCGGCGCATAGATCAGCAATTGGGC 701	

OY	361	AAAGGTCATCGAATACCTTAAACGTGGCGCTTTGGCCACTCATGGGGTACATCCCTGTGCTG	420
DB	702	AGAGTCATCGAATACCTTAAACGTGGCGCTTTGGCCACTCATGGGGTACATCCCTGTGCTG	761
OY	421	GGCGCCCCGCTCGCGCGCGCTCGCCAGAGCTCTCGGCATGGCGTGAGAGTCTTGAGAGAC	480
DB	762	GGCGCCCCGCTGGAGGGCTCGCCAGAGCTCTGGCACA CGGTGTAGGGTCTTGAGAGAC	821
OY	481	GGGGTAAATTTTGGAAACGGGAACCTTACCGGTTGCTCTTTTCTATCTTTCTTGCTGGCC	540
DB	822	GGGATAAATTAAGCAACAGGAATTTTACCGGTTGCTCTTTTCTATCTTTTGGTTGCTCT	881
OY	541	CTGCTGTCTTCATCAACACCCCGGCTCTCCGCTCGCAAGTGAAGAACATCAAGTACCGGC	600
DB	882	CTTCTGTATGTGTACATATGCTAGTCTGTGCAAGTGAAGTACAGAAACATTAAGTCTTAGC	941
OY	601	TACATGTGACTAACGACTGCACCAATGACAGATTAACCTGGCAGCTCACAGCTGTCTC	660
DB	942	TACTATCGGCACATAAATGATTGTCTAAACACAGCATCACTGGCAGCTCATGACGCAGTT	1001
OY	661	CTCCACGCTCCCGGGGTGGTCCCGTGGCAGAAAGTGGGGAAATGCATCTCACTGTGAGTA	720
DB	1002	CTTCATCTTCTCGTAAATGCTGCCATGTAGAATTAAGGCACTTGCTCTTGGCTGAGTA	1061
OY	721	CCGGTCTCAACGAAATGTGGCGGTGCGAGCGCGCCGCGGCTTCAACGACAGGGCTTGGCAGC	780
DB	1062	CAAGTAAACCCBRACGTGGCTGTGMAAACCGCGGTGGCTCATCTGTAGCTGTGMAACA	1121
OY	781	CACATCGACATATGTTGTATATGTCCGCCACGCTGTCTGTCTGTCCCTTCACTGTGGGGACCTC	840
DB	1122	CAAGTCGACATATGTATATGACAGCTACCGGCTGTCTGTCTGTATATGTGGAGATGTG	1181
OY	841	TGCGGTGGGGTGAATGCTCGACAGCCCAATATGTTCAATGTCTCGCGCAGACACA CTGAGTTT	900
DB	1182	TGCGGGGCGGTGATATATCTATCGAGGCTTTCATGTATACCAACACGCAACAATCTC	1241
OY	901	GTCCAAGCTGCATTTGCTCATCTTACCCTGGTACCATCTACTGAGACACGAGATGAGTGG	960
DB	1242	ACCMAAGTGCACACTGTTTCATCTACCAAGTGCATCACCGGCATCGATGGACATGG	1301
OY	961	GACATGATGATGAACCTGTGTCCGCCACCGGCTACACATGATCTTGGGTAGCGATGCTGTCTC	1020
DB	1302	GACATGATGCTTARCTGTGTCTTCCAACTCTTCCATGATCTTCGCTTACGCTGTGCTT	1361
OY	1021	CCCGAGGCTCAATTAATGACATCAATTAAGCGGGCTCATTTGGGGCGTATATTTGCGCTTGGCC	1080
DB	1362	CCCGAGCTGTGTCTGGAATATATTTTGGCGGCATTTGGGGTGTGGTGTGCTTGGGSC	1421
OY	1081	TACTTCTCTATGCAAGGAGCGTGGCGGAAAGTCTGTGTCTTCTGTGGCGCGCGGG	1140
DB	1422	TATTTCTTCATGACARGAGCGTGGGCCCAAGTCTRTGTGCATCTCTCTTCTTGTGGCGGA	1481
OY	1141	GTGAGCGCGGCAACCATACTGTGTTGGGGGTTCTGGCGCGAGACACACGGGGCGCTACCC	1200
DB	1482	GTGATGCGACCACTTATTTCCACGCGTACGASAGACGGGTCTGTCCGTATKMGKGGMTKCT	1541
OY	1201	AGCTTATTTGACATGGGCCCCAGCGCAGAAAAATCCAGCTCGTTAAACCAATGAGACGTGG	1260
DB	1542	RGCTCTTATTAATCTGTGTGCCAAGAGAACCTCTATTTATCAACCAATGAGACGTGG	1601
OY	1261	CACATCAACCGCACCGGCTGTAACTGTGAATGTATCTCTTGACACCGGCTTATGCGGCT	1320
DB	1602	CACATCAACCGGACGCGCTTCAATTTGCAATACAGCTATASAGACGGGTTTGMTGCTTCC	1661
OY	1321	CTGTTCTACACCCACAGCTTCAACTGTCAAGATGTCCCGAAGCATGTCCGCTGCGCGC	1380
DB	1662	YTGKTTTATCMWCCBRAGTTTCAACAGCTCTGGCTGGCCCCGAGCGCTTCTTCCGCGCC	1721
OY	1381	AGTATCGAGGCTTCCGGGTGGATATGGGCGCTTGGCAATATAGAGATTAATGTACCAAT	1440
DB	1722	GGGCTGGACGATTTTTCGATCGGCTGGGGGAACCTTGGAAATACGAAACCAAGCTACCAAC	1781

QY	1441	CAAGAGATATGAAACCTTATTTGGTCGACATCAACCAAGAGGAGTGGCGTGTCTCC	1500
Db	1782	GATBAGCAATGAGGCGGTACTGTGGCAATTAACCCCAAGGCTTGGCGCATGTGCCG	1841
QY	1501	GCGAAGACTGTGTGTGGCCGAGTGTACTGTGTTCAACCCCAAGCCAGTGTAGTGGCAGC	1560
Db	1842	GCTAGGACGGTTTGTGGACCGGCTCTATTGTATTCAACCCCTTAGCCCTGTGTGTGGGCAAC	1901
QY	1561	ACCGACAGGCTTGGAGGCGCCACTTACACGTGGGGGGAGAAATGACAGATGTCTTCTTA	1620
Db	1902	ACTACAAAGCAGGGCGTACCCACTTACACCTGGGGGRRMAAACGAGACGATGTCTTCTCG	1961
QY	1621	TTGAAACGACCTGACACACCGCTGGGGGTCAATGGTTGGGCTGCACGTGGAGTGAATCTTCT	1680
Db	1962	CTTATATGACAAAGACCCCCCGCAGAGGCTTGGTTGGGCTGCACATGTGAATGAACGGGACT	2021
QY	1681	GGCTACACCAAGACTTGTCCGGCGCACACACCTGTCCGTACTAGAGCTGACTTCAACGCAGC	1740
Db	2022	GGGTTCACTAAGACATCGGTGTGACCAACCTTGCACATTGGAAGGCTTACMACAGACACT	2081
QY	1741	ACGGAACCTGTTGTGCCCCACGGAATGTTTAAAGAGATCTGTATACACTTACCTTCAAA	1800
Db	2082	CTCATTTATTTGTGCCCCACAGACTGTGTTTAAAGAGACCAAGATGTACTTATCTTTAAG	2141
QY	1801	TGCGGCTCTGGGGCCCTTGGCTCAACGCCAAAGTGTCTGATGCACTACCCCTACAGGCTCTGG	1860
Db	2142	TGTGGAGCAGGGCCCTTGTGGTTAACTCCCAAGGTGTCTGTGTAACTAACCTTTATATATGTGG	2201
QY	1861	CATTACCCCTGTGACAGTTAACTATACCACTTCAAAAATAAGATGTATGTGGAGGGGCTT	1920
Db	2202	CATTATCCGTGCATCTGTAACTTCAACATCTTTAAGGCGCGGATGTATGTAGAGAGGGGTG	2261
QY	1921	GAGCACAAGGCTCAACGGCTGTGATGCATGCAATTTTCACTGTGTGGGATCGTTGCAACTTGGAGAC	1980
Db	2262	GAGCATGATTTCTCCGAGCATGGAACTTCAACGGCGGGAATGTGTGCAATCGAATGAAAGAT	2321
QY	1981	AGAGACAGAAAGTCACTGTCTCTCTTTTGTGCACTCCACACGGAATGGGCACTTTTAAGCT	2040
Db	2322	AGGATATGGGGGTGACGAGAGTCCACTGTCTCACTTCCACTGAATGGGGGAGTGTCTCCA	2381
QY	2041	TGCTCTTACTCGGACCGCCGCGCTTGTGACGTGTCTTCCACCTCCACCAAAACATC	2100
Db	2382	TGCTCTTCTCTGACCTACAGACCTATCCACTGGGCTTATGTGACCTTCCACCAAAACATC	2441
QY	2101	GTGACCTTACATTCATGTATGTGCTTATCACTGTGCCCTCAAAATATCATGTCCGATGG	2160
Db	2442	GTGACCTGTGACATCTTATAGGACTTTCTCCGGCTGTGCAAGATATCATGTGAAGTGG	2501
QY	2161	GAGTGGGTAACTCTTATTTCTGTCTTATGACGAGCGCCAGGGTGTGGCGCTGCTTATGG	2220
Db	2502	GAGTGGGTAACTCTCTTTCTTGTGTGTGGCAGACGCCAGGRTCTGTGATCCTTTTGG	2561
QY	2221	ATGTCTATCTTGTGTGGGCCAGGCCGGAAGACACTAGAGAAAGTGTGATCTTTGACAGCT	2280
Db	2562	ATGTCTATCAATCTGTGGCCAAAGCCGAGCGGCTGTGAAGAGTCAATCATCTTGCACTCC	2621
QY	2281	GCGAGCGAGCTAGCTCAATGCGCTTCAATTTTGTCACTTTTTCGTGGCTGTGG	2340
Db	2622	GCTAGGTGTGTATGTGCAATGTGTCCGCTGTGGTTTTTCACTTTCTTTCACAGGGGCTGG	2681
QY	2341	TACATCAAGGGGTGGGTAGTCCCTTATAGCTATTTCCCTCACTGGCGTGTGTCTTT	2400
Db	2682	TACTTAAAGGGCAGGGGTGTCCCGGTGGCAGAGTCTGTBTCTGGCTTGTGTCTTTC	2741
QY	2401	AGCTTACTGCTCTTAGCATTTGCCCAACAGGCTTATGTCTTATGACGATCTGTGCAATGCG	2460
Db	2742	CTCTCTCTAATCTCGGCTTATACACAGCAGGCTTATGCTTGTGACCGCTGTGTAAACAGGG	2801
QY	2461	CAGATAGAGGGGCTGTGCTGTAAATGATCACTCTTTACTCTCAACCCCGGGGTAAAG	2520
Db	2802	GAACTGGGGGCTGGCATAATTAATTAATTAATTCATCTTTACTTTTACCTTACCCAGCATACAG	2861
QY	2521	ACCCCTTCTCAGCGGGTTTTTGTGTGTGTGTGTACTTCTTGCACCTGGGGGAGCTAAG	2580

```
Db 2862 ATCTCCGAGCCGTTGAGTGTGGTGTCTTACATGCTGTGCTTGGCCGAGCCAG 2921
Qy 2881 GTCCAGAGAGTGGGACCAACCTATGACATGGCCGTCCTGATGATCATATATGGGCC 2640
Db 2922 ATTGACGATAGTGTCCCTCCCTGAGGTCCGAGGGGGCGTGAAGGATATCTGGAGT 2981
Qy 2641 GTCCCATATTTCTACCCAGGTGTGTGTTGACATTAACAAGAGCTCTTGGCCGAGTCTT 2700
Db 2982 GCTGATCTTACACCCACGCTTGTTGTTGAGGTACAGAAATGGTTGTTAGCAATCTG 3041
Qy 2701 GGGCTGTGTTACCTCTTAAAGTGTGTTGACGGCGTGCCTGATCTTGTGAGGGCTCAC 2760
Db 3042 GGGCTGTGTTACCTCTTAAAGTGTGTTGACGGCGTGCCTGATCTTGTGAGGGCTCAC 3101
Qy 2761 GCTGATGAGGATGTGACATGAGCAAGGATCTGCGGGGGGAGTACGTCAAGT 2820
Db 3102 GCTTGTCTACAGATGTGTACCTGTGTGAAACACCTCGGGGGGCTAGGATCATCAAGT 3161
Qy 2821 GCGCTACTAGCCCTTGGCAGGTGTGACTGGCACTTATGACCACTTACCCCTATG 2880
Db 3162 CTGTTTATACCATAGGAGATGACCGGCACTTACATGACACACCTCTCCCTTGA 3221
Qy 2881 TCGATTTGGCTGCTAGTGGCTTGGGGGACCTGGGGGTGCGCTTGAAGCTATCATCTTC 2940
Db 3222 TCAACTTGGGGCGCCAGGGTGTTCGGGACCTGGCAATCGCCGTGAGCCGTGTGTTC 3281
Qy 2941 AGTCGATGAGAAAGAAAGTATGTCGTGGGGAGCGGAGCAGCGTGTGGGGACATT 3000
Db 3282 AGCCCAATGAGAAAGTATGTCGTGGGGAGCTGAGACATGGCGGTGGAGACATC 3341
Qy 3001 TTACACGAGCTTCCCGTGTCCGCCGACTTGTGCGAGAGTCTCTTGGCCGACTGAT 3060
Db 3342 CTGATGAGCTTCCCGGTCTCCGCCAGGTAGTAGGAGTTCGTGCTGGCCGAG 3401
Qy 3061 GGCATACCTCCAAAGGGGTGAGTCTTCCGCCCATCATCTGCTTACGCCAGACACA 3120
Db 3402 GGCATACCTCCAAAGGGGTGAGTCTTACCTCCATTTACCTCCTTACCTCACCAAACT 3461
Qy 3121 CGTGGCCCTTTTGGGACATAGTGTGGTGAAGTGAACGGGGGGCGACAAGACAGACAGCT 3180
Db 3462 CGTGTCTCTGGAGTGTATGTGTGAGCTTAAACGGGGCGCGACAAAATATGACAGCT 3521
Qy 3181 GGGGAAATTCAGTGTCTGTGCAAGTCACTCAAGTCTTCTCGGAAATCCATCTCGAG 3240
Db 3522 GGGGAGTCCAGGTTCTGTCTCCGTCACAAAATTTCTTGGGGAATTCATTTCCGGC 3581
Qy 3241 GTTTTGTGAGTGTCTTACATGAGAGTGGCAAAAGACTCTGGCCGCTCAAGGGTCCG 3300
Db 3582 GTCTCTGAGACATATATCAAGGGGTGTGTATTAAGACTTGGCGGGCCCAAGGAGCA 3641
Qy 3301 GTCCGAGAGATGTATCTCAAGTGTGAGGGGACTTAAAGGGTGGCCCAAGCCCTCGG 3360
Db 3642 GTCACTCAAGATGTATCAAGGCGAGAGAGGAACTCTGTGGAAATGACCTTATGTCCTCCGG 3701
Qy 3361 ACTAAATCTTTGAGCCGCTGACAGTGTGAGCGGTGCACTGTACTTGTGTCACGCGAAC 3420
Db 3702 ACTAAGTCAATGGACCCCTGTACCTGGGGGCGGTAGACTCTTACTGTGTACCCGAAAC 3761
Qy 3421 GCTGATGTATCTCCGGCTCGAAGACGCGGGGCAAAAGGGGAGGCTATCTTCCCGAGA 3480
Db 3762 GCTGATGTATCTCCGGCTCGAAGACGCGGGGCAAAAGGATGACCGAGGGGTGATTAATCTTCCCAAG 3821
Qy 3481 CCTCTTTCACCTTGAAGGGGCTCAAGAGGCGGGTATGATGCCCCAGGGGCAAGCT 3540
Db 3822 CCTCTTCAACCTTGAAGGATCATCTCGAAGGCGCGTGTCTGTCTCAAGGGGACAGCC 3881
Qy 3541 GTCCGAGTCTTCCGGGAGCTGTGTCTCTCGGGCGTGGCTAAGTCCATGATTTATC 3600
Db 3882 GTGGGCTTGTTCAGAGCGGGGTGTGTGTGCAAGGGGTGTAGCAATCATTTGACTTATC 3941
Qy 3601 CCGGTGAGACATGTGATGTGACGGGTGCCCACTTTTATGTGACAAAGCACACCA 3660
Db 3942 CCGGTGAAATCACTGATRTCCGACACGAGCGCCCACTTTCTGTGACAAAGTTCGCCG 4001
Qy 3661 CCGTGTGAGCCGACGACTATAGTGTGGGTACTTGTGATGCCCGCATGGAGGAAG 3720
Db 4002 CAGCTGTGCCCCAGCTTATACAGTGTGTATTTGACGACCAAGGACAGCGGAAAG 4061
Qy 3721 AGCACAAGATTCCTGTGCAATATGTGCTCAGGGGTATTAAGTGTAGTGTCTTAATCC 3780
Db 4062 AGCACAAGATTCCTGCCCCGATGTGCAAGGGGTATTAAGTATCTGTACTTAATCC 4121
Qy 3781 TCAAGTGTGCTCAGCCCTGGGGGTGGGGGTGTACTTGTCTTAAGGACATAGGATCCC 3840
Db 4122 TCTGTGCGGCGCACCTGTGTGTGGGCTCTACATGTCCAAAGCCCAAGGATCAACCT 4181
Qy 3841 AACATTAGACTGTAGTGTGACACTGTGACGACCGGGGCGCCCATCAGTACTCAGATAT 3900
Db 4182 AATATGAACTGTAGTGTGAGCCGTGTACACCGGGGACTGTATCTTACTTCACTTAT 4241
Qy 3901 GGCAAATTCCTGCGCATGGGGCTGTGGCGGGCGGCTTACGACATCATATGTGAT 3960
Db 4242 GGCAGTTTATGCAATGAGGCTGTGACGCGGTGCTTATGACATCATATGTGAC 4301
Qy 3961 GAATGCCATGCGGTGACTGTACACATCTTGGATGGAACAGTCTTGTATCAAGA 4020
Db 4302 GAATGCCATTCAGTGTGACGCTACTACATCTTGGATGGAACAGTCTTGTATCAAGCT 4361
Qy 4021 GAGACAGTGGGGTGTGACTTAACTGTGTGCTTACAGTACGCTTGGGTCAGTGACA 4080
Db 4362 GAGACCGCAGGCGTCAAGGCTAGTGTGTGGCCACAGCCACGCTCCCGGTACGGTACA 4421
Qy 4081 ACCCCCAACCCCAATGAGAGAGTGGCCCTTGGGACAGAGGCGAGATCCCTTTTAT 4140
Db 4422 ACTCCCAAGTAACTAGAGAGTGGCCCTTGGTGTGACAGAGGCGAGATCCCTTTTAT 4481
Qy 4141 GGGAGGGCATTCCCTGTCTTATCATCAAGGAGAGAAAGATCTGATCTTGGCATTA 4200
Db 4482 GGGAAAGCTATTCCTTACTTGTATCAAGGGGGGAGACATCTGATCTTGTCCATTTA 4541
Qy 4201 AAGAAAAAGTGTACAGAGCTCGGGCGGCGCTTGGGGTATGGGCTTGAATCAAGTGA 4260
Db 4542 AAGAAAGATGTGACAGAGCTCGAGCGGCGCTTCCGGGGCAAGGTGTCAATGCCGTTGA 4601
Qy 4261 TACTACAGAGGTTGAGCGTCTTCGTAATACCACTCAAGGAGACGTAGTGTCTGCC 4320
Db 4602 TACTATAGGGGTCTCGACGCTCCGTTATACCACTCAAGGAGACGTAGTGTCTGCC 4661
Qy 4321 ACCGAGCCCTCATGACAGGTATATCGGGGACTTTGACTCCGTATGACGTGCAAGTA 4380
Db 4662 ACTGATGCCCTTAATGACGTGGTACACCGGCGACTTTGACTGTGTATGATGTAT 4721
Qy 4381 GCGGTCACTCAAGTTGTAGACTTCAAGTTTGAACCCCAATTCATTAACCAACAGATT 4440
Db 4722 GCAGTCTCTCAAGTTGTATCTTCAAGCTTGAACCCCACTTCAACATCACTCAAAAC 4781
Qy 4441 GTCCCTCAAGACGCTGTCTCAAGTACCAAGCGCGGGGTGCGACGGGTAGGGAGACTG 4500
Db 4782 GTCCCTCAAGACGCTGTCTCCGTAATCAAGTGAAGGAGAACTGGAGGGGGCAATTG 4841
Qy 4501 GGCATTTATAGTATTTTCACTGTGTGAGCAGCCTCAGGAATGTTTATCAGTGTAGT 4560
Db 4842 GGGRTTTCAGGATGTGTGTGTGAGGGGCGGTGTGGAGTGTGTGACACAGGTATGT 4901
Qy 4561 CTCTGTGAGTGTACAGACGAGGGGCGGATGTGTATGAGCTCAACATCGAGACACC 4620
Db 4902 CTCTGTGAGTGTATATGCGGGGAGCGTGTATGAGCTTATCACTGTGTGAGACTAG 4961
Qy 4621 GTCAGGCTCAAGGCTATTTCAACAGCGCCGCTTGTGTGCAAGACATCTTTGAG 4680
Db 4962 GTGAGACTCGGGCTATTTCAACAGCGCGGTTTCCGTAATGTCAAGACCACTGGAG 5021
Qy 4681 TTTTGGAGGAGTTTTCACCGGCTTACACATATAGTGTCCCACTTCTTCCCAAAAC 4740
Db 5022 TTCTGGGAAGCGGTCTTTTCAAGGTGTCAACATTAACGCTTCTCTTCCAGAGC 5081
```

QY	4741	AAGCAATCGGGGAAAAATTTGGCACTACTTAAACGCTTACAGGCTACAGTGTGGCTAGG	4800
Db	5082	AAGCAAGAGAGAAAACTTTGGCTATCTAAACGGCTTACAGGGCCACAGTATGGCCAGG	5141
QY	4801	GCCAAAGCCCCCCCCCTCCTCGGAGCCTCATGTTGGAAGTGTGTTGACTCGACTCAAGCCC	4868
Db	5142	GCAAAAGCCCCCTCCTTGGTGGAGCGTGAATGTGAAGTGTCTTAACTAAGGCTCAAACT	5201
QY	4861	ACACTCGTGGGCCCCCAACCTCTCTCTGAACCGCTTGGGTGCGCTGACCAAGAGTATAC	4920
Db	5202	ACACTGACTGTGTCCACCCCCCTCTGTACCGCTTGGGTGCGCTGACCAAGAGTATAC	5261
QY	4921	CTCAACATCCCGTGAAGAAATATATGCGCACCTGCAAGAGCGCACTTGAAGTATG	4980
Db	5262	TTGAGCGACCCCGTGAAGAAATATATGCGCACCTGCAAGAGCTGACCTTGAAGTATG	5321
QY	4981	ACAGCAATGGGCTTGTGGCAGGGGAGCTTGTGGCGCGCTGCGCGCTATTTGCTGCG	5040
Db	5322	ACAAGCTCATGGGCTCTGGCGGGGGGGGTGTCTAGCGCGCGTGTGCGACTTATCTGCTGCG	5381
QY	5041	ACGGGTGTGTTCGATCATTCGGCGCTTGTGACATTTAACGCGAGCGCTGTGGCGCG	5100
Db	5382	ACTGGCTGCATTTCCATCATTTGGCGCGCTTACCTGAATGTATCGGGTGTGTATCCCCY	5441
QY	5101	GACAAGAGGCTCTCTATGAGGCTTTTGTATGATGATGAGAGATGTGCTCTAGGGCGGCT	5160
Db	5442	GACAAGGABATCTTATATGAGGCTTTGATGAGATGAGAAATGCGCTCCAAAGCGCGC	5501
QY	5161	CTCATTTGAAGGGGGCAGCGGATATAGCGAGATGCTGAATTCGAAGTCCAAAGCTTATTTG	5220
Db	5502	CTCATTTGAAGAGGGCAGCGGATATAGCGAGATGCTCAATTTAAGATTCAAAGGCTCTCTA	5561
QY	5221	CAGAAAGCTTCCAAACAACTCAAGACATAACCCACTGTGCAAGGCTTACATGGCCCAAG	5280
Db	5562	CAABAGGCCACAAGGCAGAGCTCAAGRATRTAGCGACGTAACGTAATCAATGCGCCAAAG	5621
QY	5281	GTAGAACAATTTCTGGGCCCAACACATGTGAACTTTCATTAGCGGCAATCCATTAATCTCGCA	5340
Db	5622	CTTGAACAATTTTGGGCCAAACACATGTGAACTTTCATCAATGATATCAATGATATCAATCA	5681
QY	5341	GGACTATATCAACATGCGCCAGGGAGACCTGTGAGTACTTTCATGATAGGGCTTACATGCGCGC	5400
Db	5682	GGACTCTTCACCCCTACCGGGAAATCTGTGACATGTGCATCAAGATGAGCTTTTAAAGCGCGCG	5741
QY	5401	CTCAACAGTCCGCTGTCAACAAAGACACATATCTTCAACATTTTGGGGGGCTGGGCTA	5460
Db	5742	CTGACTAGGCCCATACCCACACGACACCAATCTCTTTGAACATCATGTGAGAGATGCTTG	5801
QY	5461	GCATCCCAATTTGACACACCCCGGGGGGCCACTGAGCTTCTGTTGCAATGAGGCTTATGAGGA	5520
Db	5802	GCCCTCYCAGATTTGCCCCCTCGCGGAGCCACAGGCTTCTGTTGCAATGAGGCTTATGAGGA	5861
QY	5521	GCTGCGCGAGGCAATATAGGCTTTAGGTATGAGTGTCTAGTGAACATCTGCGACGGATATGCT	5580
Db	5862	GCGGCGCTCGAAGCATATGAGCTCTGGGTATGATCTGTGAGCACTTTTGGCGGGGTATGCGC	5921
QY	5581	GCGGGCAATTTGCGGGGCTTCTGTGCGCATTCAAATTCATGTCTGCGAGAAAGCCCTCCATG	5640
Db	5922	GCAGGCAATTTCAAGGGGCGCTGTAGCTTTTAAATTCATGAGCGCGCGAAGGCCACCGTA	5981
QY	5641	GAGGATTCGTGCACTTGTGCGCTGTGGAATTTGTGCTCCGGGTGCTTGTGTATGTGAGTGC	5700
Db	5982	GAAAGCTGTGAATCTCTGCTGCTGTATATCTGCTCTGTGATGTGATGTATGTGAGTGC	6041
QY	5701	ATCTGCGCGGCATTTCTGCGCGCGACAGTGTGGAACGGGGGAAAGGCGCGCTCAATGTAGTG	5760
Db	6042	ATCTGTGAGCAATTTCTGCGCGCGCACAGTGTGCGTCAAGGGAGAGGGAGCGGTCAATGTAGTG	6101
QY	5761	AATAGACTCATTTGCTTTGCTTCCAGAGAAATCACTGTCGCCCCCACTACCTAATAGTGAC	5820
Db	6102	AACGACTCATGCTTCTTCTGCTTCCAGGGGAAACACGTTGGCCCTTACCCTCACTAATAGTGATG	6161

QY	5821	GAGTGGAGATGGCTGCGAGGCTGTGACCCAAACTACTTGGGCTCCCTTACATTAACAGACTG	5880
Db	6162	GAGTCTGACGCTTCAACGCGTGTBACGAGGCTCAAGTTCACTTACATTAACAGCTTA	6221
QY	5881	CTCAGAAAGCTCCACAACCTGATTACTGAGGACTGCGCCATGCCGCGGCTCGTGG	5940
Db	6222	CTTAGGAGACTCAATGCCCTGAGTCACTGAAGAATGGCCCACTTCCTGCGGGTCTTGG	6281
QY	5941	CTCCCCGAGTGTGGGACTGGGTTTGGACATCTCTAAGACCTTTAAATTTGGCTGACC	6000
Db	6282	CTCCAGAGACTTTGGGATTTGGTTGTCTCATCTCAAGACTTAAAACTGGCTCTCT	6341
QY	6001	TCCAATTAATTCCCAAAGATGCCGCGCTCCCTTTGTCTCTGTCAAAAAGGGGTACAG	6060
Db	6342	TCAAAATTAATCTCCCAAGATGCCGCGCATTCCTTTATCTCTTGGCAGAAAGGATACAG	6401
QY	6061	GCGCTGTGGGCGCGGCACTGGCATATGACACACAGGTGTCTTGGCGCCCAATATCTCT	6120
Db	6402	GGTGTATGGGCTGGTACGGGTTGTCTGATGACACTGTGCTCCCAAGTGGAGCAATCTCG	6461
QY	6121	GGCAATGTCCGCTGGGCTCCATGAGAAATCAAGGGGCTTAAGACTGATGAATATCTGG	6180
Db	6462	GGCCATGTCCGATGGGACATATGAATAATACAGGCCCAAGACTTGCTTGAACCTGTGG	6521
QY	6181	CAGGGGACTTTCCTATCAATTGTTACACGAGGGGCAAGTGGCTGCCAAACCCTGCGCA	6240
Db	6522	CAGGGGACTTTCCTCATTAATTGTTTACACAGAAAGGCGCTGCGTGCMAAAACCCCTCTCT	6581
QY	6241	AACCTTAAGTGGCCATCTGAGAGGGTGGGCGCTCAAGATACGGGAGGAGACAGCAGCAC	6300
Db	6582	AATTAACAAGACCGCAATTTGGAGGGTGGAGGCTGGAAATGCTTAAGATCAACAGCAT	6641
QY	6301	GGGTCTATACCACTATACATACAGAGACTCACACTGATTAATCTGAAAGTCCCTGCCAATA	6360
Db	6642	GGCTCTTCTGTGATGTAACRGGGTTAACAATGACAAACCTTAAGGTYCTTGGCCAGTGA	6701
QY	6361	CCCTCTCCGAGTCTTTTCTCTGGGTGACCGGATGCAATTCATATAGTTTGGCCCCACA	6420
Db	6702	CCAGCTCCAGAAATTTTCTCTTGGGTGACCGGGGTGCAAAATCCACGATTCGCCCCGTW	6761
QY	6421	CCGAAGCCGTTTTCGGGATGAGGCTCGTTCTGGGTTGAGGCTTAATCATTTGTGTGCTG	6480
Db	6762	CCAGGTCCCTTCTTTCCGGATGAGTAACTTCAACCTTAAGGCTTAACTCTTCTGTGCTG	6821
QY	6481	GGGTCCAGCTTCTTTCGCGACCTGAAACCGGACACAGACGTATGATGTTCATGCTAACAA	6540
Db	6822	GGCTCTCAGCTCCCTTTCGATCTCTAAGCCGAGACCGAAGGTACTGGCCTCTATGTGACA	6881
QY	6541	GATCATCTCATATACCGCGGAGACTGAGGCGGCGGCTTAAAGCGGGGGTCAACCCCA	6600
Db	6882	GACCCGTCCACATACCGCKAGGCGGAGCGAGCGATGGCAAGAGGAGATTCCTCCCY	6941
QY	6601	TCCGAGGCAAGCTCTTCGCGAGCGAGACTATCGGACACATCGCTGCGAGCACTTGACC	6660
Db	6942	TCACAGGCTAGTCTCTCAAGCGAGCAGACTCTGCGCCGCTCTGAAAGGTACTCTGACC	7001
QY	6661	ACCCACGGCAAGCCTATGATGTGACATGATGTGATGCTAACCTGTTCAATGGGGGCGAT	6720
Db	7002	ACCCATTAAGACGACATATGATGTGACATGATGTGATGCTTAACCTTTTCAATGGGAGGAT	7061
QY	6721	GTAATCTCGAATGAGTCTGGGCTCAAAAGTGCTGTTCTGAGACTCTTCGACCAATGTGC	6780
Db	7062	GTAATCCGATTTGATCTTGACCTTAAGGTGATGCTTTTGAATCTCCCTCGATTCGATGACT	7121
QY	6781	GAAAGAGAGAGCACTTGAGCTTGTGATACATCAAGATTAACATGCTCCCAAGAAAGG	6840
Db	7122	GAGGTAGAGATGATCGTAGGCTTCTGTACCATCAAGTACTGTATCAAGAGGAAG	7181
QY	6841	TTCCACACAGCTTAAACCGGCTCGGSCACAGGCTGTGATTAACACCAACCGCTTGGAGATCG	6900
Db	7182	TTCCCAACGGGCTGCTCTCTTGGGCGCGTCCAGACTACAAATCTGTTTGTGATCGAGACA	7241
QY	6901	TGAAAGAGCCAGATTAACACGGGCACTGTGGGAGCTGTGCTCTCCCTCTCTAGG	6960

```

Db      7242  TGAAGAGGCGGAGCTATGAACCAACCACTGTCTTAAGGCTGTGGCTCTCCGCCACACACT 7301
Qy      6361  AAAACCCCGAGCCTTCCGCCAAGAGAGCGCCGAGCACTGGGCTTAATGTAGACTCATTA 7020
Db      7302  CAAACGCGAGTGTCTCACTCGAGAGGCGCGCAAAATCTCGAGCCCGAGARATGTG 7361
Qy      7021  GGAATGCGCCCTTCAACAGCTGGCCATTAACTCTTTGGCCAGCCCCCAAGCGCGAT 7080
Db      7362  GAGGGGATTCCTCAGAGGAGATGGCTGACAAAGTATCAGCCCTCTCAAGAACAAATGAC 7421
Qy      7081  TCAAGCCCTTTCACAGGAGGCGGCGCTGCCGATTCGGCACTCAGAGCCTCTGATGAG 7140
Db      7422  TCCGGTCACTCCACTGAGCGGATACCGAGAGAGACATCTGTCAGCAACCTCTGAGAG 7481
Qy      7141  TTGGCCCTTTGAGAGACAGGTTCCATCTCTTCATGCCCCCCTCGAGGGGAGCTTGA 7200
Db      7482  ACTGCCCTTCAAGAACGAGGAGTCACTGTCTCTCATGCTCCCTTTCAGAGGAGACCGGA 7541
Qy      7201  GATCCAGACTTGAAGCCTGAGCAGGTAGAGCCCAACCCGCCCGAGGGGGGTGCA 7260
Db      7542  GACCCYGACTGGAGTTTGAACCAAGTGGATCCGCTCCCTTCTGAGGGGAGTGTAG 7601
Qy      7261  GCTCCCGGCTCGACTCGGGGTCTGTCTACTTGTCTCGAGAGAGACGATCCGTGCTG 7320
Db      7602  GTCAATTGTCGAGCTTAAGTCGTGTGCCACAGTCTCTGATCAAGAGGATTTCTGTATC 7380
Qy      7321  TGCCTGCTCCATGTATATCTCTTGAACCGGGGCTTAATTAATCTCTTGTATGCTCGAAG 7440
Db      7662  TGCCTCTATGTATATCTCTTGAACCGGGGCTCTAATACATATGTGTGAGGAGGAG 7721
Qy      7381  GAGAAGTTACCGATTAAACCCCTTGAAGCACTCCCTGTTCGATATCAACAAGGTATAC 7440
Db      7722  GAGAAGTTACCGATTAAACCCCTTGAAGCACTCCCTGTTCGATATCAACAAGGTATAC 7440
Qy      7441  TGTACCAACAACAAGAGCGCTCACTAAGGCTTAAGGCTTAAGGCTTAAGGCTTAAGGCT 7500
Db      7782  TCCACAACCTCGAGAGAGTGCCTCTCTGAGGCAAGAGGAGGAGCTTTTGAACAGGCTG 7841
Qy      7501  GTGCTCGACTCTCTACTAGACTCAAGTCTTAAGGCAATTAAGCTTAAGGCTTAAGGCT 7560
Db      7842  GTGCTCGAGCACTATGACTCAAGTCTTAAGGCAATTAAGGCTTAAGGCTTAAGGCT 7901
Qy      7561  ACCGCAAGGCTCTCAACCATGAGAGGCTTGCAGTTAACCCCACTTCTGCAAGA 7620
Db      7902  RGTGGAGGCTCTCAAGTGAAGAGGCTTGCAGGCTGAGCCCGCCCACTCCGCCAA 7961
Qy      7621  TCTAATAATGAGGTTGGGCTTAAGAGGCTCGAGCTTTCGGAGGGGCGTTAACCA 7680
Db      7962  TCGCGATACGGAATTTGGGCAAAAGAGGTGCGCAGCTTATCAGAGGGGCGTTAACCA 8021
Qy      7681  ATCAAGTCCGATGGAAGGAGCTCTGAGAGCTCAGAAACACCAATTCACCAACATT 7740
Db      8022  ATCCGCTCGGTGAGAGGAGCTCTCTGAGAGGCAACCACTACCCCAATGACACAACTATC 8081
Qy      7741  ATGGCCAAAATGAGGTGTCTGCTGAGACCCCAACAAGGGGGCAAGAAAGCACTGCG 7800
Db      8082  ATGGCTAAAATGAGGTGTCTGCTGAGATTCACACTAARGTGGGAAAAACCAAGCTGCG 8141
Qy      7801  CTATGCTTTAACCTGACCTTGGGCTCAAGGCTTTCGAGAGAGATGGCCCTTTATGACAT 7860
Db      8142  CTATGCTTTAACCTGACCTTGGGCTCAAGGCTTTCGAGAGAGATGGCCCTTTATGACAT 8201
Qy      7861  ACACAAAAATCTTCCAGGCGGTGATGGGGCTTCTTATGATTCAGATATTCGCCGCT 7920
Db      8202  KCACAAAAATCTTCCAGGCGGTGATGGGGCTTCTTATGATTCAGATATTCGCCGCT 8261
Qy      7921  CAGCGGTGAGGTTCTCTTGAAGCATGGCGGAAAGAGAGACCTATGGGTTTTTCG 7980
Db      8262  GAGCGGTGAGGTTCTCTCTCAAGCTTGGGGAAGTAAAGAGACCAATGGGGTTCTCG 8321
Qy      7981  TATGATACCGGATGCTTTGACTCAACCGTCACTGAGAGAGCATGAGAGTGC 8040

```

```

Db      8322  TATGACACCCGCTGCTTTGACTCAACCGTCAAGAGAGGAGCATTAAGAACAGAGATCC 8381
Qy      8041  ATATATGAGGCGCTGCTCTTTCGCCAGAGAGGCCCACTGGCATCACTGCTTAATGAG 8100
Db      8382  ATATATGAGGCTTGTCTTCTGCTCAAGAGGCAAGACGTATATACCTGCTCACTGAG 8441
Qy      8101  AGACTTTACGTGGAGGAGCTTATGTTCAACAGCAAGGCAAACTTGCAGGTACAGGCT 8160
Db      8442  AGACTTTACGTGGAGGAGCTTATGTTCAACAGCAAGGCAAACTTGCAGGTACAGGCT 8501
Qy      8161  TGCCTGCGAGGAGGAGTGTCTCAACACTAGCAATGGGGAACACATATCAATGATCGTAA 8220
Db      8502  TGCCTGCGAGGAGGAGTGTCTCAACACTAGCAATGGGGAATCAATGATGATGATCAAT 8561
Qy      8221  GCTTATGAGGAGTGTAAAGCTCAGAGATTAATGCGCCCAATGCTGTATGCGCGAT 8280
Db      8562  GCTTATGAGGAGTGTAAAGCTCAGAGATTAATGCGCCCAATGCTGTATGCGCGAT 8621
Qy      8281  GACTTGTGTATCTCAGAAAGCCAGAGGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAG 8340
Db      8622  GACTGTGTGTATCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8681
Qy      8341  TTACAGAGAGGATATGACAGAGATTTCTGCCCCCTCTGCTGAGACCCGCCAGACGAGT 8400
Db      8682  TTACAGAGAGGATATGACAGAGATTTCTGCCCCCTCTGCTGAGACCCGCCAGAGAT 8741
Qy      8401  GATCTGAGCTGATTAATCTTGTCTCTCAAAATGCTGTGAGGCTGAGGAGAGAGAG 8460
Db      8742  GACTTGTGTATCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8801
Qy      8461  CGCGGAGATTAATCTTGTCTCTCAAAATGCTGTGAGGCTGAGGAGAGAGAGAGAG 8520
Db      8802  CGCGGAGATTAATCTTGTCTCTCAAAATGCTGTGAGGCTGAGGAGAGAGAGAGAG 8861
Qy      8521  ACAGTTAAGACATCCCTCTGATTAATGAGTGTGAGGAGAGAGAGAGAGAGAGAGAG 8580
Db      8862  ACAGTTAAGACATCCCTCTGATTAATGAGTGTGAGGAGAGAGAGAGAGAGAGAGAG 8921
Qy      8581  ATATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8640
Db      8922  ATATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8981
Qy      8641  CTGAGCAAGACCTCAACTTTGAGATGATGATGATGATGATGATGATGATGATGATG 8700
Db      8982  CTGAGCAAGACCTCAACTTTGAGATGATGATGATGATGATGATGATGATGATGATG 9041
Qy      8701  CTGAGCAAGACCTCAACTTTGAGATGATGATGATGATGATGATGATGATGATGATG 8760
Db      9042  CTGAGCAAGACCTCAACTTTGAGATGATGATGATGATGATGATGATGATGATGATG 9101
Qy      8761  CCCCAAGAACTGACAGGAGCTTCAAGCTTCAAGAACTTCAAGAACTTCAAGAACTT 8820
Db      9102  CCCCAAGAACTGACAGGAGCTTCAAGCTTCAAGAACTTCAAGAACTTCAAGAACTT 9161
Qy      8821  GCGTGAAGAGCGGAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 8880
Db      9162  GCGTGAAGAGCGGAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 9221
Qy      8881  GCGTGAAGAGCGGAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 8940
Db      9222  GCGTGAAGAGCGGAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 9281
Qy      8941  TTGCGGAGAGCGGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 9000
Db      9282  TTGCGGAGAGCGGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 9341
Qy      9001  GACATTTATCAAGGCTGTGCGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 9060
Db      9342  GACATTTATCAAGGCTGTGCGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 9401
Qy      9061  CTTTGTGAGGAGTGTGCGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 9099
Db      9402  CTTGCGTGAAGTGTGCGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 9440

```

RESULT 7
AAC86645
ID AAC86645 standard; DNA; 9611 BP.
XX
XX AAC86645;
AC
XX
XX 02-APR-2001 (first entry)
DT
DE Nucleotide sequence of chimeric Hepatitis C virus clone pH77CV-J65.
XX
XX HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine; ss.
XX
OS Synthetic.
OS Hepatitis C virus.
FH
FH Key Location/Qualifiers
FT CDS 342..9389
FT /*tag= a
XX
XX MO200075338-A2.
XX
XX 14-DEC-2000.
XX
XX 02-JUN-2000; 2000MO-US015446.
XX
XX 04-JUN-1999; 99US-0137693P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Vanagi M, Bukh J, Emerson SU, Purcell RH;
XX
XX WPI; 2001-061728/07.
XX
XX Nucleic acid molecule encoding human hepatitis C virus of genotype 2a for
PT developing vaccines, for diagnosis of hepatitis C virus and in screening
PT assays for identification of antiviral agents.
XX
XX Disclosure; Page 100-103; 167P; English.
XX
XX AAC86645-48 represent chimeric cDNA clones of infectious Hepatitis C
CC virus (HCV). In each clone the C, E1 and E2 genes are derived from HCV
CC strain HC-J6CH genotype 2a. The p7 protein was derived from HCV strain HC
CC -J6CH or HCV strain pcv-H77C, and the NS genes were derived from HCV
CC strain pcv-H77C. Such HCV sequences are capable of expressing the virus
CC when transfected into cells. The HCV protein is useful for assaying
CC candidate antiviral agents for activity against HCV. Antibodies specific
CC for HCV polypeptide are useful in prevention and treatment of diseases
CC caused by HCV in animals, in particular humans. The HCV polypeptides
CC serve as immunogens in the development of vaccines for preventing HCV in
CC mammals or as antigens in diagnostic assays for detecting the presence of
CC HCV in biological samples. The HCV polynucleotide is also useful for
CC identifying cell lines capable of supporting the replication of HCV in
CC vitro and to produce attenuated viral strains via passage in vitro or in
CC vivo
XX
XX Sequence 9611 BP; 1935 A; 2896 C; 2682 G; 2098 T; 0 U; 0 Other;
SQ
Query Match 59.3%; Score 5399.2; DB 5; Length 9611;
Best Local Similarity 75.0%; Pred. No. 0;
Matches 682; Conservative 0; Mismatches 2203; Indels 78; Gaps 3;
QY 1 ATGAGCAAAATCTTAACCTCAAGAAAAACCAAAACCAACCAACGTCGCCCA 60
DB 342 ATAGGCAAAATCTTAACCTCAAGAAAAACCAAAACCAACCAACGTCGCCCA 401
QY 61 GAGGTTAAGTTCCGGGCGCGCGAGATGTTGGCGGAGATATCTTTGGCGCGAGG 120
DB 402 GAGGTTAAGTTCCGGGCGCGCGAGATGTTGGCGGAGATATCTTTGGCGCGAGG 461
QY 121 GGGCCGAGGTTGGGTGTGCGCGGCAAGAGAACTTCGAGCGGTCACGACGCTGA 180
|||||

DB 462 GGGCCGAGGTTGGGTGTGCGCGGCAAGAGAACTTCGAGCGGTCACGACGCTGA 521
QY 181 AGGCGGCAAGCCATTCCTTAAGATCGGCGTCCATCTGACAAATCTTGGGAAAAACAGGA 240
DB 522 AGGCGGCAAGCCATTCCTTAAGATCGGCGTCCATCTGACAAATCTTGGGAAAAACAGGA 581
QY 241 TACCCCTGGCCCTTATACGGGAATGAGGGACTCGGCTGGGACAGATAGCTCTGTCCCC 300
DB 582 TACCCCTGGCCCTTATACGGGAATGAGGGACTCGGCTGGGACAGATAGCTCTGTCCCC 641
QY 301 CGAGGTTCCCGTCTTGGGGCCCCCAATGACCCCGGAGTATGTCGCAACGTCGAGT 360
DB 642 CGAGGTTCCCGTCTTGGGGCCCCCAATGACCCCGGAGTATGTCGCAACGTCGAGT 701
QY 361 AAGGTATGATATACCTTAAGTATGTCGCTTGGCGGATATGAGGATATCCCTGTGTGT 420
DB 702 AAGGTATGATATACCTTAAGTATGTCGCTTGGCGGATATGAGGATATCCCTGTGTGT 761
QY 421 GGGCGCCGCTCGGCGCGCTCGGCAAGAGCTTCGCGCATGGCGTGAAGATCTTGAAGAC 480
DB 762 GGGCGCCGCTCGGCGCGCTCGGCAAGAGCTTCGCGCATGGCGTGAAGATCTTGAAGAC 821
QY 481 GGGGTTAATTTTGAACAGGAACTTAACCGGTTGCTCTTTTCTATCTTTCTTGGGCC 540
DB 822 GGGGTTAATTTTGAACAGGAACTTAACCGGTTGCTCTTTTCTATCTTTCTTGGGCC 881
QY 541 CTGCTGTCTGATATCAACCCCGGCTCCGCTGCGGAGTGAAGAACATACATACCGGC 600
DB 882 CTGCTGTCTGATATCAACCCCGGCTCCGCTGCGGAGTGAAGAACATACATACCGGC 941
QY 601 TACATGATGATTAACGACTGCAACCAATGACAGATTAACGAGCTCCAGGCTGTGTC 660
DB 942 TACATGATGATTAACGACTGCAACCAATGACAGATTAACGAGCTCCAGGCTGTGTC 1001
QY 661 CTCACGTCCTCCGGGTCGTCCTCGTGAGAAAGTGGGAAATGATCTAGTCGTGATA 720
DB 1002 CTCACGTCCTCCGGGTCGTCCTCGTGAGAAAGTGGGAAATGATCTAGTCGTGATA 1061
QY 721 CCGGTCACCGAATGTGGCGGTGACAGGCCCCGGGCTTACAGCAGAGGCTTGGGACG 780
DB 1062 CCGGTCACCGAATGTGGCGGTGACAGGCCCCGGGCTTACAGCAGAGGCTTGGGACG 1121
QY 781 CACATGCAATGATGATGATGTCGCGCAAGCTGTGCTGCTGCTTACAGTGGGAGACTC 840
DB 1122 CACATGCAATGATGATGATGTCGCGCAAGCTGTGCTGCTGCTTACAGTGGGAGACTC 1181
QY 841 TCGGTTGGGTTGATGCTCGGAGCCCAATGTTGATTTCTGCGCGACCACTGGTTT 900
DB 1182 TCGGTTGGGTTGATGCTCGGAGCCCAATGTTGATTTCTGCGCGACCACTGGTTT 1241
QY 901 GTCCAAAGACTGCAATGCTTCATCTTACCTGTGATCACTGATCAACCGGATGGCATGG 960
DB 1242 GTCCAAAGACTGCAATGCTTCATCTTACCTGTGATCACTGATCAACCGGATGGCATGG 1301
QY 961 GACATGATGATGAATGATGTCGCGCAAGGCTTACCATGATCTTGGCGTACGAGATGCTGTC 1020
DB 1302 GACATGATGATGAATGATGTCGCGCAAGGCTTACCATGATCTTGGCGTACGAGATGCTGTC 1361
QY 1021 CCGAGGTCATTAATGACATCAATTAACGAGGCTTATGAGGCGTCAATGTTGGCTTGGCC 1080
DB 1362 CCGAGGTCATTAATGACATCAATTAACGAGGCTTATGAGGCGTCAATGTTGGCTTGGCC 1421
QY 1081 TACTTCTCTATGAGGAGGATGAGGAAAGTGTGATCTCTTGTGTTGGCGCGCGG 1140
DB 1422 TACTTCTCTATGAGGAGGATGAGGAAAGTGTGATCTCTTGTGTTGGCGCGCGG 1481
QY 1141 GTGAGCGCGGCAACCAATCTGTTGGGAGTTCTGCGCGGACCAACCGGCGCTTACCC 1200
DB 1482 GTGAGCGCGGCAACCAATCTGTTGGGAGTTCTGCGCGGACCAACCGGCGCTTACCC 1541
QY 1201 AGCTTATTTGACATGGGCCCCAGGACGAAATTCAGCTGTTTAAACCAATGGCAGCTGG 1260
DB 1542 AGCTTATTTGACATGGGCCCCAGGACGAAATTCAGCTGTTTAAACCAATGGCAGCTGG 1601

QY 1261 CACATCAACGCGACCGGCTGGAATGACTCTTGGACACCGGCTTTATCGGCTT 1320
DB 1602 CACATCAACGCGACCGGCTGGAATGACTCTTGGACACCGGCTTTATCGGCTT 1661
QY 1321 CTGTTCTACACCGACGCTTCAACTCGTCAGAGATGTCGCGAAGCATGTCGCGTCCG 1380
DB 1662 CTGTTCTACACCGACGCTTCAACTCGTCAGAGATGTCGCGAAGCATGTCGCGTCCG 1721
QY 1381 AGTATGAGGCGCTTCGCGGTGGAGATGGGGGCGCTTGAATATAGAGATATGTCACCAAT 1440
DB 1722 AGTATGAGGCGCTTCGCGGTGGAGATGGGGGCGCTTGAATATAGAGATATGTCACCAAT 1781
QY 1441 CCAGAGATATAGAACCTTATTTGCTGGCACTACCCCAAGGCGAGTGGCGTGTCTCC 1500
DB 1782 CCAGAGATATAGAACCTTATTTGCTGGCACTACCCCAAGGCGAGTGGCGTGTCTCC 1841
QY 1501 GCGAAGACTGTGTGGCGGCGAGTGTACTGTTTCAACCCCGAGCGGATGTTGGGCAAG 1560
DB 1842 GCGAAGACTGTGTGGCGGCGAGTGTACTGTTTCAACCCCGAGCGGATGTTGGGCAAG 1901
QY 1561 ACCGACAGGCTTGGAGGCGCCACTTACACGTTGGGGGAGAAATGAGACAGATGTCCTCTA 1620
DB 1902 ACCGACAGGCTTGGAGGCGCCACTTACACGTTGGGGGAGAAATGAGACAGATGTCCTCTA 1961
QY 1621 TTGAACAGCACTCGACCAACGCGTGGGGTCAATGTTGGGCTGCAAGTGAATGAATCTTCT 1680
DB 1962 TTGAACAGCACTCGACCAACGCGTGGGGTCAATGTTGGGCTGCAAGTGAATGAATCTTCT 2021
QY 1681 GGGCTACACCAAGACTTGGCGGCGACACACCTGCGTACTAGAGCTGACTTCAACGCCAGC 1740
DB 2022 GGGCTACACCAAGACTTGGCGGCGACACACCTGCGTACTAGAGCTGACTTCAACGCCAGC 2081
QY 1741 ACGGACCGTGTGTGGCGGCGGAGCGTGTGTTAGGAAGCATCTGATACACTTACTCTGAA 1800
DB 2082 ACGGACCGTGTGTGGCGGCGGAGCGTGTGTTAGGAAGCATCTGATACACTTACTCTGAA 2141
QY 1801 TGGGCTCTGGGCGCTGCTGACGCGCAAGGTGCTGATGCACTACCCCTACAGGCTCTGG 1860
DB 2142 TGGGCTCTGGGCGCTGCTGACGCGCAAGGTGCTGATGCACTACCCCTACAGGCTCTGG 2201
QY 1861 CATTAACCTCTGACAGTTAACTATACATCTTCAAAATAGAGATGTTGGGAGGGGTT 1920
DB 2202 CATTAACCTCTGACAGTTAACTATACATCTTCAAAATAGAGATGTTGGGAGGGGTT 2261
QY 1921 GAGCAGAGGCTCAAGGCTGACATGCAATTCTGCTGGGGAGATCGTTGCAATTGAGAGAC 1980
DB 2262 GAGCAGAGGCTCAAGGCTGACATGCAATTCTGCTGGGGAGATCGTTGCAATTGAGAGAC 2321
QY 1981 AGAGACAGAGTCAACTGTCTCTCTTGTGTCATCTCCACCAAGGATGGGCAATTTTACT 2040
DB 2322 AGAGACAGAGTCAACTGTCTCTCTTGTGTCATCTCCACCAAGGATGGGCAATTTTACT 2381
QY 2041 TGCTCTTACTGGAGCTGCGCGCTTGTGCACTGCTCTTCTCCACTCTCCACCAAAACATC 2100
DB 2382 TGCTCTTACTGGAGCTGCGCGCTTGTGCACTGCTCTTCTCCACTCTCCACCAAAACATC 2441
QY 2101 GTGACGTAACAATTCATGATAGGACTATACCTGCGCTCACAAAATACATGTCGAGATG 2160
DB 2442 GTGACGTAACAATTCATGATAGGACTATACCTGCGCTCACAAAATACATGTCGAGATG 2501
QY 2161 GAGTGGGTAAATCTTTATTCCTGCTCTTGAAGGACGCGCAGGGTGTGGCGCTGTATGG 2220
DB 2502 GAGTGGGTAAATCTTTATTCCTGCTCTTGAAGGACGCGCAGGGTGTGGCGCTGTATGG 2561
QY 2221 ATGCTCATCTTGTGGGCGAGCGGACGAGACGTAAGAAAGCTGTATCTTGAAGCT 2280
DB 2562 ATGCTCATCTTGTGGGCGAGCGGACGAGACGTAAGAAAGCTGTATCTTGAAGCT 2621
QY 2281 GCGAGCGAGCTAGCTGCAATGGCTTCCATATTTTGTGCACTTTTGTGGCTGCTGG 2340
DB 2622 GCGAGCGAGCTAGCTGCAATGGCTTCCATATTTTGTGCACTTTTGTGGCTGCTGG 2681

QY 2341 TACATCAAGGTCGAGTACTCCCTTAGTACTATTCCTCACTGAGGCTGTGTCTTT 2400
DB 2682 TACATCAAGGTCGAGTACTCCCTTAGTACTATTCCTCACTGAGGCTGTGTCTTT 2741
QY 2401 AGCTTACTGCTCTTAGACTTTGCCCAACAGGCTTATGCTTATGAGCATCTGTGATGGC 2460
DB 2742 AGCTTACTGCTCTTAGACTTTGCCCAACAGGCTTATGAGCATCTGTGATGGC 2801
QY 2461 CAGATAGAGCGGCTGCTGGTAAATGATCACTCTTATCTCTCACCCCGGATTAAG 2520
DB 2802 TGTGTGGCGGGCTTGTCTTGTGGGTTAAATGGCGCTGACTGTGTGGCATATTAAG 2861
QY 2521 ACCCTTCAAGCGGTTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
DB 2862 CCGTATATCAGCTGGTGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2921
QY 2581 GTCCAGGATTTGGGACCAACCTATGACAGGTCGCGGTGGCGGATGGGATGATGGGCG 2640
DB 2922 GTCCAGGATTTGGGACCAACCTATGACAGGTCGCGGTGGCGGATGGGATGATGGGCG 2981
QY 2641 GTCCAGATATTTCTAACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2700
DB 2982 ATGT 3041
QY 2701 GGGCTCTGTACTCTCTTAAAGGTCTTGAACGCGGTGCTGTGTGTGTGTGTGTGTGTGT 2760
DB 3042 GGGCTCTGTACTCTCTTAAAGGTCTTGAACGCGGTGCTGTGTGTGTGTGTGTGTGTGT 3101
QY 2761 GCTCTACTAGATGTGACACATGTGCAAGGCAATTCGCGGGGGGCGAGTACGTCAGATG 2820
DB 3102 GGGCTCTACTAGATGTGACACATGTGCAAGGCAATTCGCGGGGGGCGAGTACGTCAGATG 3161
QY 2821 GCGCTACTAGGCTTGGCAGTGTGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2880
DB 3162 GCGCTACTAGGCTTGGCAGTGTGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3221
QY 2881 TCGAATGGGCTGCTAGTGGGCTTGGGAGACTTGGCGGTGCGGTGATCATCTTC 2940
DB 3222 TCGAATGGGCTGCTAGTGGGCTTGGGAGACTTGGCGGTGCGGTGATCATCTTC 3281
QY 2941 AGTCCGATGGAAGAAATGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3000
DB 3282 AGTCCGATGGAAGAAATGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3341
QY 3001 TTACAGGACTTCCGCTGTCCGCGGCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3060
DB 3342 TTACAGGACTTCCGCTGTCCGCGGCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3401
QY 3061 GGGCTATATCTCAAGGAGT 3120
DB 3402 GGGCTATATCTCAAGGAGT 3461
QY 3121 CGTGGCGTTTTGGGCAACATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3180
DB 3462 CGTGGCGTTTTGGGCAACATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3521
QY 3181 GGGGAAATTCAGGCTCTGT 3240
DB 3522 GGGGAAATTCAGGCTCTGT 3581
QY 3241 GTTTTGTGACTGTCTTACATGAGGCTGTGCAACAAGCTGTGCGGCTTCAAGGGTCCG 3300
DB 3582 GTTTTGTGACTGTCTTACATGAGGCTGTGCAACAAGCTGTGCGGCTTCAAGGGTCCG 3541
QY 3301 GTTACGAGATGTACTCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3360
DB 3642 GTTACGAGATGTACTCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3701
QY 3361 ACTTAAATCTTTGGAGCGGT 3420
DB 3702 TCCGCGCTATTTGACACCGCTGTACTGTGGGCTCTTCCGAGCTTTTACGAGTGTGTGT 3761
QY 3421 GCTGATGTCAATCCCGGCTGGAAGACGCGGGGACAAACGCGGAGGCTTACTCTCCCGAGA 3480

Db 3762 GCGATGTCATTCGCCGCGCGCGGAGGTATAGCAGGGGTACCTGCTTCCCGCCG 3821
Qy 3481 CTTCTTTTCACCTTTGAAGGGGTCTCAAGAGAGCCCGGTGCTATGCCAGGGGCAAGCT 3540
Db 3822 CCCATTTCTCACTTTGAAGGGGTCTCTCGGGGGGTCCGCTGTGTGTCGCCCGGGGACAGGCC 3881
Qy 3541 GTCCAGGCTTTCCGGGAGAGCTGTGTGCTCGGGGCGGTAGTCAATAGATTTTCATC 3600
Db 3882 GTGGGCTATTTCAAGGCGCGCGGTGTGACCCGTGAAGTGGCTTAAGCGGTGACCTTTATTC 3941
Qy 3501 CCCGTGAGACACTCGACATCGTCACGCGGTCCCACTTTAGTGAACAAGACACCA 3660
Db 3942 CTGTGTGAGAACTTAGGGGACAAACATAGATCCCGGTGTTCACGAGCAACTCTCTCCA 4001
Qy 3661 CTGTCTGTGCCCCAGACCTATCAGGTGGGTACTTTGATCGCCGACCTGGCAGTGGAAAG 3720
Db 4002 CCAGCACTGTGCCCCAGACCTTCCAGGTGGCCCACTGTGATGCTCCACCGGCAAGCGTTAAG 4061
Qy 3721 AGCAACCAAGTTCCTGTGCGATATGCTCAAGGGGTATTAAGTGTAGTCTTAATCCC 3780
Db 4062 AGCAACCAAGTTCCTGTGCGATATGCTCAAGGGGTATTAAGTGTAGTCTTAATCCC 4121
Qy 3781 TCAGTGCTGCAACCTGTGGGGTTTGGGCGTACTTGTCTAAGGCACTGGCATCAATCCC 3840
Db 4122 TCTGTGTCTGCAACGCTGGGCTTTGTGTGCTTGAATGTCAATGTCCAAAGGCCCATGGGGTTGATCT 4181
Qy 3841 AACATTAGACGTGAGTCAAGACTGTGACGACCGGGGCGCCCATCAGTACTCCACATAT 3900
Db 4182 AATATCTAGGACCGGGGTGAGAAACAATTACACTGGACGCCCATCAGTACTCCACCTAC 4241
Qy 3901 GGGAAATTCCTGCGCGATGCGGGGCTGTGCGGGCGCGCCCTACGACATCATATATGTGAT 3960
Db 4242 GGGAAATTCCTGCGCGATGCGGGGCTGTGCGGGGCTGTGCGGGGCTGTGCGGGGCTGTGCG 4301
Qy 3961 GAATGCCATGCGCGTGAACCTTACCAACCATCTTTGGATCGGAAAGTCTTGAATCAAGA 4020
Db 4302 GAGTGCCACTCCAGGATGCGACATCATCTTGGGCACTCGGCACTGTCTTGAACCAAGA 4361
Qy 4021 GAGACAGCTGGGTGAGACTTACGTGTGTGCTGCTACAGCTACGCCCTGGGTGCTAGACA 4080
Db 4362 GAGACTGCGGGGGGAGACTGTGTGTGTGCTGCTACCTCCCGGGCTCGGTACT 4421
Qy 4081 ACCCCCCACCCCAATAGAGAGGTGGCCCTTGGGAGAGAGGCGAGATCCCTTGTAT 4140
Db 4422 GTGTCCATCTTAACATGAGAGGTGTCTGTCTCACACCGGAGAGATCCCTTGTAC 4481
Qy 4141 GGGAGGCGATTCCTCTGTCTTACATCAAGGAGAGAAAGATCTGTCTGCAATTTCA 4200
Db 4482 GGGAGGCTATCCCTCTCGAGGTGATCAAGGGGGGAAAGATCTGTCTGTGCACTCA 4541
Qy 4201 AAGAAAAAGTGTGACGAGCTCGCGGGCGCCCTTGGGGGTATGGGCTTGAATCTAGTGGCA 4260
Db 4542 AAGAAAAAGTGTGACGAGCTCGCGGGAGCTGGTTCGATTTGGGCAATATGCGGTGGCC 4601
Qy 4261 TACTACGAGAGGTGTGACGCTCTCCGTATACCACTCAAGGAGCGTAAAGTGTGTGCGC 4320
Db 4602 TACTACGAGAGGTGTGACGCTCTCCGTATACCACTCAAGGAGCGTAAAGTGTGTGCGC 4661
Qy 4321 ACCGACGCGCTCACTGACAGGGTATACCTGGGAACTTTGACTCCGTGATCGACTGCAAGSTA 4380
Db 4662 ACCGATGCTCTCACTGACAGGGTATACCGGGGAACTTTGACTCTGTGATAGACTGCAAGCG 4721
Qy 4381 GCGGTCACTCAAGTTGTGACCTTCAAGTTTGAAGCCCACTTACCATTAACCAACAGATT 4440
Db 4722 TGTGTCACTCAAGCTGTGATTTGAGCTTGAAGCTTCACTTATAGAACCAACAGCG 4781
Qy 4441 GTCCCTTAAGAGCTGTCTCAAGTAAAGCGAGCGCGGGGTGCAAGGGTAAAGGGAAGACTG 4500
Db 4782 CTCCCGCAAGATGTGTCTCAAGACTCAAGCGCGGGGCAAGACTTGGCAAGGGGAAGCCA 4841
Qy 4501 GGCATTTATAGATATGTTTTCACCTGTGAGCGAGCTCAGGAATGTTTGAACAGTGTAGTG 4560
Db 4842 GGCATTTATAGATATGTTTTCACCGGGGAGCGCCCTCCGGGCAATGTTGACTGTGCTGTC 4901
Qy 4561 CTCTGTAGTGTCTACAGCGAGGGGCGGCAATGTTATAGTCAACCATATGAGAACCAAC 4620
Db 4902 CTCTGTAGTGTCTACAGCGGGGCTGTGCTTGTGTATAGTCAAGCGCCCGGAGACTACA 4961
Qy 4621 GTCCAGGCTCAAGGGGCTATTTCAACAGCGCGGGTGTGCTGTGTGACCAAGCAATCTTGAAG 4680
Db 4962 GTTATGCTTAGAGCTAGATGAAACACCCCGGGGCTTCCGTGTGCGAGACATCTTGA 5021
Qy 4681 TTTTGGAGGCGAGTTTTCACCGGCTTCAACACATAGATGCCCACTTCTTTCCCAACA 4740
Db 5022 TTTTGGAGGCGAGCTTTCACGGGCTTCACTCATATAGATGCCCACTTTCCTTCAACA 5081
Qy 4741 AAGCAATCGGGGGGAAAAATTTCCGATCTTAAACAGCCCTACAGCTTAAGTGTGCGTAGG 4800
Db 5082 AAGCAAGTGGGAGAACTTTCCTTAACTGTGTAGCGTACCAAGCACGCTGTGCGTAGG 5141
Qy 4801 GCCAAAGCCCCCCCCGCTGCTGAGAGCTATGTAAGTGTGACTCGACTCAAGCCC 4860
Db 5142 GCTTCAGCCCTCTCCCATGTGTGGACCAAGTGTGAAGTGTGATTCGCCCTTAAACC 5201
Qy 4861 ACACTGTGGGCCCCACACCTCTCTGTACCGCTTGGGCTGTGTAAACAAGAGTCAAC 4920
Db 5202 ACCCTTCATGGGCCAACACCCCTGTATTAACAAGTGGGCGCTGTGAGAAATGAAGTCAAC 5261
Qy 4921 CTCAACATTCGGGTGACGAAATATATGCGCACTGTGATGCAACCGCACTTTAGGTATG 4980
Db 5262 CTGACGCAACCAATCAACAAATATCATATGATGATGATGCGCGCACTGGAGGTGCTC 5321
Qy 4981 ACCAGCAATGGGCTTGTGGGAGGGGAGTCTTGGCGGCGCTGCGCGGCTATTTGCGTGGC 5040
Db 5322 ACAGCACTGGGTGTCTGTGTGGCGGCTCTGTGGCTCTGCGCGGCTATTTGCGTGTCA 5381
Qy 5041 ACCGGGTGTGTTCATCATATGCGCGCTTGCACATTAACCAAGCGAGCGTGTGCGCGC 5100
Db 5382 ACAGGCTGTGTGTATGATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5441
Qy 5101 GACAAAGAGTCTCTTATGAGGCTTTTGTATGATGAGATGAGAAATGTGCTTATAGGCGGCT 5160
Db 5442 GACAAAGAGTCTCTTATGAGGCTTTTGTATGAGATGAGAAATGTGCTTATAGGCGGCT 5501
Qy 5161 CTCATTGAAAGAGGGGAGCGGATAGCGAGATCTGAAAGTCCAAAGATCCAAAGCTTATTTG 5220
Db 5502 TACATGAGCAAGAGATGATGTCTGCTGAGCACTTCAAGGAGAAAGGCTCTGCGCTCTG 5561
Qy 5221 CAGCAAGCTTCAAAACAGCTCAAGATTAACCAACCACTGTGAGAGGCTTATAGGCCCAAG 5280
Db 5562 CAGACGCGCTCCGCGCATGAGAGGTATACCCCTGTGTCCAGAACCACTGGCAGAA 5621
Qy 5281 GTAGAACATTTCTGGGCAAAACATGTGGAATCTTATAGCGGATCCAAATACCTGCGCA 5340
Db 5622 CTGAGAGTCTTTTGGGCAAGACATGTGGAATTTTATCATGCTGGGATACAAATCTTGGCG 5681
Qy 5341 GAGCTATCAACACTGCGAGGGAACCTGCAAGTGTCTTCAATGATGAGGCTTCAAGTCCGCC 5400
Db 5682 GAGCTGTCAACGCTGCTGTGTAAACCCGCAATGCTTCAATGATGAGGCTTTCATACAGCTGCC 5741
Qy 5401 CTCAACAGTCCGCTGTCAACCAACCACTATCTTCTCAACATTTTGGGGGCTGTGCTA 5460
Db 5742 GTCAACAGGCCCACTTAACCACTGCGCAAAACCTCTCTTCAACATATTTGGGGGCTGTGCT 5801
Qy 5461 GCATCCCAATTTGACACACCCGCGGGGCACTGTGCTGTGTGTGTGTGTGTGTGTGTGTGT 5520
Db 5802 GTGCGCAAGTCTGCGCGCCCGGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5861
Qy 5521 GTGCGGATGAGTATAGGCTTATAGGCTTATAGGCTTATAGGCTTATAGGCTTATAGGCTT 5580
Db 5862 GCGCCATTCGAGCGGTGTGACTGTGGGAGGTCTGTGTGACATTTCTTGTGAGGATATGCG 5921
Qy 5581 GCGGGGCAATTTTGGGGGCTCTGTGTGCAATTCAGATCATATGTCTGTGAGAAAGCCCTTCAG 5640
Db 5922 GCGGGGCTGTGCGGAGCTTGTGTGATTCAGATCATATGAGCGGTATGCTCCCTCCACG 5981

QY 5641 GAGGATGTCGCAACTTGGCTGCTGGAATTCGTCTCGGGTGCCTTGGTAGTGGAGTC 5700
DB 5982 GAGGACCTGTCATCTGCTGCCCCGCATCTCTGCGCTGGAGGCTTTGTAGTGGGTG 6041
QY 5701 ATCTGCGGCGCATTTCTGCGCGGACGCTGGGACCGGGGAAAGCGCGTTCATATGATG 5760
DB 6042 GTCTGCGGCAAAATCTGCGCGGACGCTTGGCGCGGCGGAGGAGGAGTGCATATGATG 6101
QY 5761 AATGACTCATTTGCTTGTCTTCCAGAGAAATCAAGTGGCGCCCGACCACTACGATG 5820
DB 6102 AACCGGCTAATAGCTTGTCTTCCGCGGAAACATGTTTCCCGACGACTACGTCG 6161
QY 5821 GAGTCGATGCGTGGCAGCGGTGACCCCACTATTTGGCTCCCTTACCATTAACGAGCTG 5880
DB 6162 GAGAGCGATGAGCGCGCGGTCACTGCATACCTACAGCAGCTCACTGTAACCAAGTC 6221
QY 5881 CTGAGAACTCCACAACTGGATTACTGAGAGCTGCCCATCCCATGCGGGGCTGCTG 5940
DB 6222 CTGAGGCGACTGCACTAGTGAATGCTGAGAGTGTACACTGCATGCTCGGTTCTG 6281
QY 5941 CTGCGGATGTCGAGGAGCTGGGTTTGACATCTCAACAGACTTTAAATAATGGCTGACC 6000
DB 6282 CTAAAGGACATCTGGGACTGGATATGCGAGGTGCTGAGCGACTTTAAGACTTGGCTGAAA 6341
QY 6001 TCCAAATTAATCCCAAGATGCCGCGCTCCCTTTGTCTGCTGTCAAAAGGGTACAG 6060
DB 6342 GCAGAGCTCATGSCCAACTGCTGGGATTCCTTTGTGTCCTGCGACGCGGGTATAGG 6401
QY 6061 GCGCTGTCGCGCGGACCTGGATCATGACACAGGTGTCTTGGCGGCGCCAAATCTCT 6120
DB 6402 GGGGTCTGGCGAGGAGCGGATTAATGACACTGCTGCGCACTGTGAGAGCTGGAATCACT 6461
QY 6121 GGCATATGTCCTTGGGCTCCATGAGATCAAGGGGCTTAAGACCTGCGATGAATATCTG 6180
DB 6462 GGAATGTCAAAAACGGGACGATGAGTGTGTCGCTTAGGACTTGGAGAAACATGTG 6521
QY 6181 CAGGGGACCTTCTCTAATTTGTTTACAGGAGGGCCAGTGCCTGGCGAAACCGCGCCA 6240
DB 6522 AGTGGAGCGTTCCCATTAAGCGCTACACAGGGGCGCTGTACTCCCTTCTGCGCG 6581
QY 6241 AACTTTAAGTTCGCATCTGAGAGGTGGCGGCTCAAGGTACGGGAGTGAACCGACAC 6300
DB 6582 AACTATAAGTTTCGCGTGTGAGAGGTGTCTGCAAGAGATAGTGAATAGCGGGTG 6641
QY 6301 GGGTCATACCACTACATAACAGGACTCACACTGATTAATTGAAGTCCCTGCGCACTA 6360
DB 6642 GGGGACTTCCACTAGGTATCGGGTATGACTACTGACAACTTTAAATGCCGTGCGAGTC 6701
QY 6361 CCTCTCCGAGTTCTTTTCTGGGTGAGCGGAGTGAGATCAATAGGTTTGCCTCCCA 6420
DB 6702 CCATGCGCCGAATTTTTCAGAAATTTGAGCGGGGTGCGCTTACACAGGTTTGCCTCC 6761
QY 6421 CCGAAGCGGTTTTTCCGGAGTGAAGTCTGCTTGGCGTTAATTTCAATTTGTGTC 6480
DB 6762 TGCAAGCCCTTGTGCGGAGAGGTATCATTCAGAGTGAAGTTCAGAGTACCCGCT 6821
QY 6481 GGGTCCAGCTTCTTGGGACCGCTGAACCGGACAGAGAGTATTGATTCATGCTAAACA 6540
DB 6822 GGGTGCATAATTAATCTTGGAGCGCGAACCAGAGTGAAGTGTGTAAGTCCATGCTCACT 6881
QY 6541 GATTCATCTCATATCAAGCGGAGAGCTGACGCGCGCTTTAGCGGGGCTCAACCCCA 6600
DB 6882 GATCCCTCCCATATTAACAGAGAGGCGCGCGGAGAAAGTTGGCGAGAGGCTCAACCCCT 6941
QY 6601 TCCGAGGAGAGCTCTCGCGGAGGAGGATTCGAGACCATGCGTGGCGAGCGACCGACCC 6660
DB 6942 TCTATGCGAGTCTCTCGGCTAGCGAGTGTCCATCTCTCAAGGCAACTTGGACCC 7001
QY 6661 ACCCAGCGCAAGGCTATGATGAGCATGTGATGCTAACT-----GTTTC 6708
DB 7002 GCACACATGACTCCCTGAGCGCGAGCTCATAGAGCTTAACCTCTCTGAGAGGAGAG 7061

QY 6709 ATGGGGGGCGAGTGTGACTCGGATAGAGTCTGGGTCGAGTCCAAAGGTGCTTCTGACTCTTC 6768
DB 7062 ATGGCGGCAACATCAACAGGTTGAGTCAGAGAACAAAGGTGATTTCTGAGCTTCTTC 7121
QY 6769 GACCCAAATGTCGAAAGAAAGAGCGACTTTGAGCCTTGATACCATAGAAATATGATGTC 6828
DB 7122 GATCCGCTTGTGCAAGAGAGATGAGCGGAGTCTCGTACCTGCAAAATTTCTGCGG 7181
QY 6829 CCCAAGAAAGGTTCCACAGCTTTACCGGCTGGGACGCGCTGATTACAAACCGACCG 6888
DB 7182 AAGTCGAGAGATTCGCCCGGCGCGCTGCGCTGAGGCGGCGGAGCTTCAACCCCGCG 7241
QY 6889 CTGTGGAATCTGTGAAAAGCCGATTAACAAACGCGCACTGTGCGGCTGTGCTCTTC 6948
DB 7242 CTAGTAGAGCGTGGAAAAAGCCTGACTAGAACCACTGTGTGTCATGTGCTGCGCTGA 7301
QY 6949 CCTCCTCTAGAGAAAACCGCGAGCGCTCCCGCAAGAGGCGCGGACAGTGGGCTTAAT 7008
DB 7302 CCACCTCAAGGTCCTCTGCTGCTGCGCTCGGAAAAAGGTACGGGTGCTTCACCC 7361
QY 7009 GAGGACTCCATAGAGATGCTCCCTTCAACAGCTGCGCATTAAGTCTTTGGCGACCGCC 7068
DB 7362 GAATCAACCTTATCTACTGCTTGGCGAGCTTGGCACCAAAAGTTTGGAGCTCTCA 7421
QY 7069 CCAAGCGCGGATTCAGGCTTTTCACTGCGGCGGCGGCTGCGGATTCGCGAGTCAAGC 7128
DB 7422 ACTTCGCGGATTAACGGGCAATATACGAACAATCTCTGAGCGCGC----- 7469
QY 7129 CCTCCTGATGATGTGGCCCTTTCGAGACAGGTTCCATCTTTCAATGCGCGCGCTGAG 7188
DB 7470 CTTTCTGGCTGCCCCCGACCTCGAGCTTGAAGTCTTATTTCCATGCCCCCTCTGAG 7529
QY 7189 GGGGACTTGAAGATCCAGACTGAGCCTGAGCAGTGAAGAGCCCAACCCCCCCAG 7248
DB 7530 GGGGAGCTGGGATCCGGA----- 7549
QY 7249 GGGGGGGTGGAGCTCCCGCTGAGCTCGGGGTCTGTGTTACTTGTCTCGAGAGAGAC 7308
DB 7550 -----TCTCAGCGACGGGTATGTGTACAGGTGAGTGGGCGGACAG 7595
QY 7309 GACTCGGTGTGTCGTCTCATATGTATATCTCTGAGACCGGGGCTTAATTAATCTCTTGA 7368
DB 7596 GAAAGATGTGTGTGCTGTCAATGTCTTATTTCTTGAACAGGCGCACTGTCACCCGTG 7655
QY 7369 AGTCCGAGAGGAGAGGATTAACCGATTAACCCCTTGAAGCACTCCCTGTGGGATATCAC 7428
DB 7656 GCTGCGGAGAGAAAGAAAGTCCCATCAACGCACTGAGCACTGTTGCTAGCGCATCAC 7715
QY 7429 AACAGGTGTACTGTACACACAAAGAGCGCTCTCACTAAGGCTTAAAAAGTAACTTTT 7488
DB 7716 AATCTGTGTATTCACCACTTCACGAGGTGCTTGCAGAGGAGAAAGTCACTATT 7775
QY 7489 GATAGAGTGAAGTGTGCACTCTCACTACAGTCAAGTCTTAAGGACATTAAGCTAGG 7548
DB 7776 GACAGACTGCAAGTCTTGGAGAGCACTTACAGAGAGTGTCTCAAGAGGTCAAGAGAGG 7835
QY 7549 GCTTCAAGGTCAACCGCAAGCTCTCACCATAGAGAGGCTTGCAGTTAACCCGACCC 7608
DB 7836 GCGTCAAAAGTGAAGCTAATCTGTATCGTGAAGAGAGCTTGCAGCTGAGCGCCCA 7895
QY 7609 CATTTCTGAAGATCTTAATATGAGGTTTGGGCTTAAGAGGTCCGAGCTTGTCCGGAGG 7668
DB 7896 CATTCAGCCAAATTCAGAGTTTGGCTATGGGCAAAAGAGTCCGTTGCAATGCAAAAG 7955
QY 7669 GCGGTTAACCAATCAAGTCCGTGTGGAAGAGCTCTGGAAGGACTCAAGAAACCACAATT 7728
DB 7956 GCGGTAGCCCAATCAATCTCGTGTGAAAAGACTTCTGGAAGACAGTGTAAACCAATTA 8015
QY 7729 CCCCAACCAATTAATGCGCAAAATTAAGTGTCTGTGCGGACCCCAAGAGGGGCGCAAG 7788
DB 8016 GACACTACATATATGCGCAAGAGAGGTTTCTGCGTTCAAGCTGAGAGGGGGGTGCT 8075
QY 7789 AAAGCACTGCTTATCTTTTACCTGACCTGCGGTCAAGGCTGTGCGAAGAAATGGCC 7848

```

Db      8076 AAGCAGACGTCGTCATCGTGTCCCGACCTGGGCGCGGTGCGAGAAATGAGCC 8135
Qy      7849 CTTTATGACATTAACAAAAAATCTCTCAGCGGTGATGAGGGCTTTATGATTCAG 7908
Db      8136 CTGTAACACGTGTGTAACAAAGCTCCCTGGCGGTGATGGAAAGCTCTACGATTCAA 8195
Qy      7909 TATTCGCCGCTCAGCGGTGAGAGTTTCTCTTGAAGACATGGGCGGAAAGAAAGACCT 7968
Db      8196 TACTCACACAGACAGCGGGTTGAATTCCTGTGACAGCGTGAAGTCCAGAGAACCCCG 8255
Qy      7969 ATGGTTTTTGTATGATACCCGATGCTTGACTCAAACGTCATGAGAGACATCAG 8028
Db      8256 ATGGGTTTCTGTATGATACCCGCTGTTTGTACTCAAGTCACTGAGAGGACATCGT 8315
Qy      8029 ACTGAGAGATCCATATATCGGGCTGCTCTTGGCCGAGAGAGCCCACTGCCATAC 8088
Db      8316 ACCGAGAGGCAATTTACCAATGTTGACCTGGAACCCCAAGCCCGGTGGCCATCAG 8375
Qy      8089 TCGCTAAGTGAAGACTTTAGTGTGAGAGGCTATGTTTCAACAGCAAGGCGCAACCTGC 8148
Db      8376 TCCTTCACTGAGAGGCTTTATGTTGGGGCCCTTTACCAATTCAGAGGGGAAATCTGC 8435
Qy      8149 GGGTACAGGGGTTGCCGCGCAGCGGGGTGTCTACCACTAGCATGGGAAACCATCACA 8208
Db      8436 GGCTACGACGAGGTGGCGCGAGCGGTACTGACCACTAGCTGTGTAACACCTCACT 8495
Qy      8209 TGTACTGTAAAGCTTTAGCGGCTTTGAAGCTGACAGGATATCGGCCCAATCTGT 8268
Db      8496 TGTACTCAATCAAGGCGCGGAGCCTGTGAGCGCAGAGGCTCCAGAGCTGACCATCTC 8555
Qy      8269 GTATGCGGCGATGACTTGTGTGATCTCAGAAAGCCAGGAGGACCGAGAGAGAGACGG 8328
Db      8556 GTGTGTGCGCACACACTTATGTTATCTGTGAAGTCCGGGGGTCTGAGAGAGACGGCG 8615
Qy      8329 AACCTGAGAGCTTTCACGAGGCTATGACCAAGTATCTGCCCTCTGTGTGACCCGCC 8388
Db      8616 AGCTGAGAGCTTTCACGAGGCTATGACCAAGTATCTGCCCTCTGTGTGACCCGCC 8675
Qy      8389 AGACCGGAGTATGATCTGAGAGCTGATTAACATCTTGTCTCTCAATGTGTCTGTGCGCTG 8448
Db      8676 CAACCAAGATATCACTTGAAGCTTATTAACATCACTCTCTCAACGTGTGATGCTGCCAC 8735
Qy      8449 GGCCCAAGGCGCGCAGATCTACTGACGAGAACCTTCACTCCATCGAGCGCCG 8508
Db      8736 GAGGCGCTGAAAGAGGCTTACTTACCTTACCGTGAACCTTACACCCCTCTGCGAGA 8795
Qy      8509 GCTGCTGGGAAACAGTTAGACACTCCCTGTCAATTCATGCTGGGAAACATCATCAG 8568
Db      8796 GCGGCGTGGGAGACAGAACACACTCCAGTCAATTCCTGGGTAGGCAATATCATG 8855
Qy      8569 TAGGCCCCGACATATGGGCTCGATGTCTCTGTAGACACTTCTTCTCATTTCTCATG 8628
Db      8856 TTTGCCCCCAGCTGTGTGGGAGAGATATCTGATGACCATTTCTTACCGTCTCATATA 8915
Qy      8629 GCTCAAGACAGCGCTGACGACAACTTCACTTTGAGATGTACGAGAGGCTGTACTCGTG 8688
Db      8916 GCCAGGATCAGCTTGAACAGGCTTTTAACTGTGATCTACGAGGCTGTACTTACATA 8975
Qy      8689 AGTCCCTTGGACCTCCAGCTATATTAATTAAGTTACATGGGCTTGAACGCTTTTCTGTG 8748
Db      8976 GAACCACTGATCTACTCTCAATCATTCATAAAGACTCATGAGCCATTTTTCATC 9035
Qy      8749 CAACATATCACTCCCAAGACATGACACAGGCTGCTTACGCTTCAAAAAATTTGGGGCG 8808
Db      9036 CACAGTACTCTCAGGTGAATTAATATGAGGTGCGCATAGCTTCAAAAAATTTGGGGCTC 9095
Qy      8809 CCAACCTCTCAGAGGTGGAAGAGCGCGGACAGTGCATCAGAGGCTCTCATCTCCGT 8868
Db      9096 CCGGCTTGTGAGGTTGAGACACCGGCGCGAGGCTCCGCGTGTAGGCTTCTCTCAGA 9155
Qy      8869 GGGGAGAGAGCGGCGCTTTCGCTGATCTCTTCAATTTGGGCGGTGAAGACCAAGCTC 8928

```

```

Db      9156 GAGGACAGGCTGTATATGTGCAAGTACTCTTCACTGAGCGAGTAAGAACAAAGCTC 9215
Qy      8929 AAATCACTCATTGCGGGAAGGCGGCTCTGGAATTAACAGTGGTTCAACGCTGGC 8988
Db      9216 AAATCACTCATTGAGGCGGCGGCTGAGCTGACTGTGCGGTGTACAGGCTGGC 9275
Qy      8989 GCGGCGGCGGCGACATTTATACAGGCTGTGCGTCCGAGACCCGCTTATTTGCTTT 9048
Db      9276 TAGAGCGGGAAGACATTTATACAGGCTGTCTCATGCCCCGCGCTGTTCTGATTT 9335
Qy      9049 GGCCTACTCTTCTTTTGTAGGAGGCTTTTCTTACTCTCCCGCTCG 9098
Db      9336 TGCTACTCTGCTGCTGCGAGGAGTACGATCTCTCCCAACCG 9385

```

RESULT 8

AAC86647 standard; DNA; 9611 BP.

AAC86647;

02-APR-2001 (first entry)

Nucleotide sequence of chimeric Hepatitis C virus clone J6S.

HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine; ss.

Synthetic.

Hepatitis C virus.

Key Location/Qualifiers

FT CDS 342..9389

FT CDS /**tag= a

FN WO200075338-A2.

PD 14-DEC-2000.

PF 02-JUN-2000; 2000MO-US015446.

PR 04-JUN-1999; 99US-0137693P.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Yanagi M, Bukh J, Emerson SU, Purcell RH;

DR WPI; 2001-061728/07.

PT P-PSDB; AAB30732.

PT Nucleic acid molecule encoding human hepatitis C virus of genotype 2a for

PT assays for identification of antiviral agents.

PS Disclosure; Page 131-134; 167pp; English.

CC AAC86645-48 represent chimeric cDNA clones of infectious Hepatitis C virus (HCV). In each clone the C, E1 and E2 genes are derived from HCV strain HC-J6CH genotype 2a. The p7 protein was derived from HCV strain HC-J6CH or HCV strain pcV-H7C, and the NS genes were derived from HCV strain pcV-H7C. Such HCV sequences are capable of expressing the virus when transfected into cells. The HCV protein is useful for assaying candidate antiviral agents for activity against HCV. Antibodies specific for HCV polypeptide are useful in prevention and treatment of diseases caused by HCV in animals, in particular humans. The HCV polypeptides serve as immunogens in the development of vaccines for preventing HCV in mammals or as antigens in diagnostic assays for detecting the presence of HCV in biological samples. The HCV polynucleotide is also useful for identifying cell lines capable of supporting the replication of HCV in vitro and to produce attenuated viral strains via passage in vitro or in vivo

Sequence 9611 BP; 1935 A; 2896 C; 2682 G; 2098 T; 0 U; 0 Other;

Query Match 59.3%, Score 399.2, DB 5, Length 9611,
Best Local Similarity 75.0%, Pred. No. 0,
Matches 6829, Conservative 0, Mismatches 2203, Indels 78, Gaps 3,
QY 1 ATGAGCAAAATCTCTAAACCTCAAGAAAAACCAAAAGAAACCAACCGTCCGCCGAA 60
DB 342 ATGAGCAAAATCTCTAAACCTCAAGAAAAACCAAAAGAAACCAACCGTCCGCCGAA 401
QY 61 GACGTTAAGTTCCGGGCGGCGGCGAGATCGTTGGCGAGATATCTTTGGCCGCGAGG 120
DB 402 GACGTTAAGTTCCGGGCGGCGGCGAGATCGTTGGCGAGATATCTTTGGCCGCGAGG 461
QY 121 GGGCCCAAGTTGGGTGTGCGCGCAAGAGAACTTGGAGCGGTCCAGCCAGTGA 180
DB 462 GGGCCCAAGTTGGGTGTGCGCGCAAGAGAACTTGGAGCGGTCCAGCCAGTGA 521
QY 181 AGGGGCAAGCCATCCCTAAAGATCGGCGCTCCACTGGCAATCTGGGGAAAAACAGGA 240
DB 522 AGGGGCAAGCCATCCCTAAAGATCGGCGCTCCACTGGCAATCTGGGGAAAAACAGGA 581
QY 241 TACCCCTGAGCCCTATACGAGGAATGAGGGACTCGGCTGGGAGAGATGGCTTCCGCC 300
DB 582 TACCCCTGAGCCCTATACGAGGAATGAGGGACTCGGCTGGGAGAGATGGCTTCCGCC 641
QY 301 CGAGGTTCCGTCCTCTTGGGAGCCCAATGACCCCGGCAATGAGTGGCGCAAGTGG 360
DB 642 CGAGGTTCCGTCCTCTTGGGAGCCCAATGACCCCGGCAATGAGTGGCGCAAGTGG 701
QY 361 AAGGTCATCGATACCTCTTAACGTGGGCTTTGGCCGACCTGAGGGTATCATCCCTGTG 420
DB 702 AAGGTCATCGATACCTCTTAACGTGGGCTTTGGCCGACCTGAGGGTATCATCCCTGTG 761
QY 421 GGGCCCGCGCTGGCGGCGCTGCGCAGAGCTCTCGGCGATGCGTGAAGTCTTGAAGAC 480
DB 762 GGGCCCGCGCTGGCGGCGCTGCGCAGAGCTCTCGGCGATGCGTGAAGTCTTGAAGAC 821
QY 481 GGGGTTAATTTTGGCAACAGGAACTTACCGGATGCTCTTTTCTATCTTCTGCTGAC 540
DB 822 GGGGTTAATTTTGGCAACAGGAACTTACCGGATGCTCTTTTCTATCTTCTGCTGAC 881
QY 541 CTGCTCTCTGATACCAACCCCGGCTCCGCTCGGCGAAGTGAAGAACTACGATCCGGC 600
DB 882 CTGCTCTCTGATACCAACCCCGGCTCCGCTCGGCGAAGTGAAGAACTACGATCCGGC 941
QY 601 TACATGTGATTAACGACTGCAACATGACAGATTAACCTGGCAGCTCAGAGCTGCTGTC 660
DB 942 TACATGTGATTAACGACTGCAACATGACAGATTAACCTGGCAGCTCAGAGCTGCTGTC 1001
QY 661 CTCCAGTCCCGGGTGCCTCCGCTGGGAGAAAGTGGGAAATGCATCTCACTGCTGATA 720
DB 1002 CTCCAGTCCCGGGTGCCTCCGCTGGGAGAAAGTGGGAAATGCATCTCACTGCTGATA 1061
QY 721 CCGGCTACCGGAATGGCCGCTGAGAGGCGCCGCGGCTCAAGAGGGCTTGGCGAGCG 780
DB 1062 CCGGCTACCGGAATGGCCGCTGAGAGGCGCCGCGGCTCAAGAGGGCTTGGCGAGCG 1121
QY 781 CACATCGACATGTGTGATGTCGCGCAGGCTTGTCTGCTCTGACCTTACGTTGGGAGACCTC 840
DB 1122 CACATCGACATGTGTGATGTCGCGCAGGCTTGTCTGCTCTGACCTTACGTTGGGAGACCTC 1181
QY 841 TGGCGTGGGGTGTATGCTCGCAGGCCCAATGTTCAATGTCTGCGCGAGCAACACTGTTT 900
DB 1182 TGGCGTGGGGTGTATGCTCGCAGGCCCAATGTTCAATGTCTGCGCGAGCAACACTGTTT 1241
QY 901 GTCCAGACTGCAATTTGCTCACTCAACCTGTATACATCACTGACACCGGATGGCATGG 960
DB 1242 GTCCAGACTGCAATTTGCTCACTCAACCTGTGTACATCACTGACACCGGATGGCATGG 1301
QY 961 GACATGATGATGAATCTGTGCGCCAGGCTACATGATCTTGGCGTACGCGATCGTGTG 1020
DB 1302 GACATGATGATGAATCTGTGCGCCAGGCTACATGATCTTGGCGTACGCGATCGTGTG 1361
QY 1021 CCCGAGGTCAATTATAGACATCAATTAGCGGGGCTCAATGGGCGTCACTGCGCTTGGCC 1080

DB 1362 CCCGAGGTCAATTATAGACATCAATTAGCGGGGCTCAATGGGCGTCACTGCGCTTGGCC 1421
QY 1081 TACTTCTATGAGGAGAGCGTGGGAGAAAGTCGTTGTGATCATCTTCTGTGGCCGCGGAG 1140
DB 1422 TACTTCTATGAGGAGAGCGTGGGAGAAAGTCGTTGTGATCATCTTCTGTGGCCGCGGAG 1481
QY 1141 GTGAGCGCGGCAACCCATATCTGTGGGAGTCTGCCGCGAGAACACCGGGCGCTCAAC 1200
DB 1482 GTGAGCGCGGCAACCCATATCTGTGGGAGTCTGCCGCGAGAACACCGGGCGCTCAAC 1541
QY 1201 AGCTTATTTGACATGGGCCCCGAGGAGAAATTCAGCTGTGTTAAACCAATGGCAGCTGG 1260
DB 1542 AGCTTATTTGACATGGGCCCCGAGGAGAAATTCAGCTGTGTTAAACCAATGGCAGCTGG 1601
QY 1261 CACATCAACGCGACCGCCCTGAACCTGCAATGACTCCTTGACACACCGGCTTTATCGGCTCT 1320
DB 1602 CACATCAACGCGACCGCCCTGAACCTGCAATGACTCCTTGACACACCGGCTTTATCGGCTCT 1661
QY 1321 CTGTTCTAGACCCACAGCTTCAACTGCTGAGAGATGCCGAAAGCATGTCGCTTGC 1380
DB 1662 CTGTTCTAGACCCACAGCTTCAACTGCTGAGAGATGCCGAAAGCATGTCGCTTGC 1721
QY 1381 AGTATGAGGCTTCCGGGTGGGATGGGCGCTTGGCATATGAGGATATATGACCAAT 1440
DB 1722 AGTATGAGGCTTCCGGGTGGGATGGGCGCTTGGCATATGAGGATATATGACCAAT 1781
QY 1441 CCAGAGATATGAGACCTATTGCTGGGCACTACCCCAAGGCAAGGCGAGTGGGCTG 1500
DB 1782 CCAGAGATATGAGACCTATTGCTGGGCACTACCCCAAGGCAAGGCGAGTGGGCTG 1841
QY 1501 GCGAAGACTGTGTGTGGCCCAAGTATCTGTTTCAACCCCAAGGCAAGGCTGAGTGGGCAAG 1560
DB 1842 GCGAAGACTGTGTGTGGCCCAAGTATCTGTTTCAACCCCAAGGCTGAGTGGGCAAG 1901
QY 1561 ACCGACAGGCTTGGAGCGCCCACTTACAGTGGGGGAGAAATGAGACATGTCCTTCC 1620
DB 1902 ACCGACAGGCTTGGAGCGCCCACTTACAGTGGGGGAGAAATGAGACATGTCCTTCC 1961
QY 1621 TTGAACAGCACTGACCAACCGCTGGGGTCAATGGTCTGCGCTGACGTTGAGAACTCTTCT 1680
DB 1962 TTGAACAGCACTGACCAACCGCTGGGGTCAATGGTCTGCGCTGACGTTGAGAACTCTTCT 2021
QY 1681 GGGTACACCAAGACTTGGGCGGCAACCCCTGCGATGATGAGCTGACTTCAACGCGCAGC 1740
DB 2022 GGGTACACCAAGACTTGGGCGGCAACCCCTGCGATGATGAGCTGACTTCAACGCGCAGC 2081
QY 1741 ACGGACCTGTTGTGCCCCCAAGCATGTTTTAGGAAGCATTCGATACCTTACCTCAAA 1800
DB 2082 ACGGACCTGTTGTGCCCCCAAGCATGTTTTAGGAAGCATTCGATACCTTACCTCAAA 2141
QY 1801 TGGGCTCTGGGCGCTTGGCTCAACGCAAGGTGCTGATGACTTACCCCTACAGGCTCTGG 1860
DB 2142 TGGGCTCTGGGCGCTTGGCTCAACGCAAGGTGCTGATGACTTACCCCTACAGGCTCTGG 2201
QY 1861 CATTACCCCTGCAACAGTTAATATACATCTTCAAAATAGATGATGTGGAGGGGTT 1920
DB 2202 CATTACCCCTGCAACAGTTAATATACATCTTCAAAATAGATGATGTGGAGGGGTT 2261
QY 1921 GAGCACAGGCTCACCGGCTGATGCAATTTCACTCGTGGGATGTTGGCAACTTGGAGGAC 1980
DB 2262 GAGCACAGGCTCACCGGCTGATGCAATTTCACTCGTGGGATGTTGGCAACTTGGAGGAC 2321
QY 1981 AGAGACAGAACTGATGTCCTCTTGTGACACTCCACCAAGGATGGGCACTTTTAACT 2040
DB 2322 AGAGACAGAACTGATGTCCTCTTGTGACACTCCACCAAGGATGGGCACTTTTAACT 2381
QY 2041 TGTCTTTACTCGAGACCTGCGCGCTTGTGAGCTGTTCTTCACTTCCACCAAAATC 2100
DB 2382 TGTCTTTACTCGAGACCTGCGCGCTTGTGAGCTGTTCTTCACTTCCACCAAAATC 2441
QY 2101 GTGACGTACAAATTCATGATGAGCTTACCTGCGCTCACAAAATACATCGTCCGATGG 2160

2442 GTGACCTACATTCATGTATGCGCTATCACTGCGCTCACAAATAATCATGTCGATGG 2501
Qy
2161 GAGTGGTAAATCTTATTCCTGCTCTTACGGGACCCAGGGTTTGCGCTGCTATGG 2220
Db
2502 GAGTGGTAAATCTTATTCCTGCTCTTACGGGACCCAGGGTTTGCGCTGCTATGG 2561
Qy
2221 ATGCTCATCTTGTGTGGCCAGGCGGAGCAGACATAGAGAACTGTGCTATCTTGCAGCT 2280
Db
2562 ATGCTCATCTTGTGTGGCCAGGCGGAGCAGACATAGAGAACTGTGCTATCTTGCAGCT 2621
Qy
2281 GCGAGCGCAGCTAGCTGCAATGCTCTCTATATTTTGTGCTATCTTGTGTGCTGCTGG 2340
Db
2622 GCGAGCGCAGCTAGCTGCAATGCTCTCTATATTTTGTGCTATCTTGTGTGCTGCTGG 2681
Qy
2341 TACATCAAGGGTGGGGTAAAGTCCCTTATGCTATCTTATCTTATCTTATCTTATCTT 2400
Db
2682 TACATCAAGGGTGGGGTAAAGTCCCTTATGCTATCTTATCTTATCTTATCTTATCTT 2741
Qy
2401 AGCTTACTGCTCTAGCATTGCCCCAACAGGCTTATGACGCTATGCTATGCTATGCT 2460
Db
2742 AGCTTACTGCTCTAGCATTGCCCCAACAGGCTTATGACGCTATGCTATGCTATGCT 2801
Qy
2461 CAGATAGAGCGGCTGCTGTGTATGATCACTCTTATCTTATCTTATCTTATCTTATCT 2520
Db
2802 TCGTGTGGCGCGCTGCTGTGTATGATCACTCTTATCTTATCTTATCTTATCTTATCT 2861
Qy
2521 ACCCTTCTAGCCGCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
Db
2862 CGCTATATCAAGCTGT 2921
Qy
2581 GTCCAGAGAGTGGGACACCATATGACAGTGGCGCGGTGTGTGTGTGTGTGTGTGTGT 2640
Db
2922 CTGCACTGT 2981
Qy
2641 GTCCCATATTTCTACCCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2700
Db
2982 ATGT 3041
Qy
2701 GGGCTGT 2760
Db
3042 GGAACCCCTTTGT 3101
Qy
2761 GCTCTACTGAGATGT 2820
Db
3102 GGGCTTCTCCGATCTGT 3161
Qy
2821 GGGCTACTAGCCCTGT 2880
Db
3162 GCCATCATCAAGTTAGGGGGCTTACTGGCACTATGTATATCACTTCACTTCACTT 3221
Qy
2881 TCGGATTTGGGCTGT 2940
Db
3222 CGAGACTGT 3281
Qy
2941 AGTCCGATGAGAGAAAGTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3000
Db
3282 TCCCGAATGAGAGAAAGTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3341
Qy
3001 TTACACGAGCTTCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3060
Db
3342 ATCAACGAGCTTCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3401
Qy
3061 GGGCTATATCTTCAAGGGT 3120
Db
3402 GGAATGTCTTCAAGGGT 3461
Qy
3121 CGTGGCTTTTGGGCAACATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3180
Db
3462 AAGAGGCTCTCTAAGGT 3521
Qy
3181 GGGGAAATTAAGGTCTGT 3240
Db
3522 GGTGAGGTTCAGATGT 3581

Qy
3241 GTTTTGTGACTGTGTACATGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3300
Db
3582 GTATGTGTGACTGTGTACATGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3641
Qy
3301 GTCAAGCAGATGTGTACTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3360
Db
3642 GTATGTGAGATGTGTATCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3701
Qy
3361 ACTAATCTTTGTGAGCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3420
Db
3702 TCCCGCTCATTTGACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3761
Qy
3421 GGTGATGT 3480
Db
3762 GCGATGT 3821
Qy
3481 CTTCTTTTCACTTTGAGGGGTCTCTAGAGAGCCCGGTGTGTGTGTGTGTGTGTGTGT 3540
Db
3822 CCAATTTCTTACTTTGAAAGGCTCTCGGGGGGTGTGTGTGTGTGTGTGTGTGTGTGT 3881
Qy
3541 GTGGAGCTTTTCCGGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3600
Db
3882 GTGGGCTTATTTAGGGGCGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3941
Qy
3601 CCGTTTGTGACACTGT 3660
Db
3942 CTTGTGTGAAACCTTGT 4001
Qy
3661 CTTGT 3720
Db
4002 CAGAGAGT 4061
Qy
3721 AGACCAAGT 3780
Db
4062 AGACCAAGT 4121
Qy
3781 TCAGT 3840
Db
4122 TCTGT 4181
Qy
3841 AACATTGTGAGCTGT 3900
Db
4182 AATATGTGAGCTGT 4241
Qy
3901 GGGAAATTTCTGT 3960
Db
4242 GGGAAATTTCTGT 4301
Qy
3961 GAATGT 4020
Db
4302 GAGT 4361
Qy
4021 GAGACAGCTGT 4080
Db
4362 GAGACAGCTGT 4421
Qy
4081 ACCCCCCCACTTGT 4140
Db
4422 GT 4481
Qy
4141 GGGAGGGCGATTTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4200
Db
4482 GGGAGGGCGATTTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4541
Qy
4201 AAGAAAAAGT 4260
Db
4542 AAGAAAAAGT 4601
Qy
4261 TACTTACAGAGGTTGT 4320
Db
4602 TACTTACAGAGGTTGT 4661

Qy 4321 ACCGACGCCCTCATGACAGGGATATCTGGGACCTTGTGACTCCGATGACCTGCAAGCTA 4380
Db 4662 ACCGATGCTCATGACGCTGCTTAAACGGGACCTTGCATCTGTGATATGACTGCAACAG 4721
Qy 4381 GCGGTCACTCAAGTTTGTAGACTTCACTTACACCCACATTACCAATTAACACAGATT 4440
Db 4722 TGTGTACATCGACAGCTGATTTTCAAGCTTACCCCTACCTTTACCATTTAGACACACAG 4781
Qy 4441 GTCCCTCAAGACGCTGTCTCAAGTATAGCCAGGCGCGGGGTGCAAGGATGAGGGAAGCTG 4500
Db 4782 CTCCCCACAGATGCTGTCTTCAAGACTCAACGCCGGGGAGAGACTGGCAGGGGGAAGCCA 4841
Qy 4501 GGCATTTATAGTATGTTTCACTGTGAGCGAGCTTCAGGAATGTTTGAACATGTATGTG 4560
Db 4842 GGCATTTATGATTTGTGTGACACCGGGGAGCGCCCTCCGGACATGTTGACATCTGTCGTC 4901
Qy 4561 CTCTGTAGTGTCTCAAGACGAGGGCCGCAATGTATGACTCAACACATCGGAACACACC 4620
Db 4902 CTCTGTAGTGTCTATGACGCGGGCTGTGTCTGTATGACTCAGCGCCCGCGAAGACTACA 4961
Qy 4621 GTCAAGCTCAAGGCGTATTTTCAACACGCGCGTGTGCTGTGTGCAAGACATCTTGAG 4680
Db 4962 GTTAGGCTACAGAGCTATATGACACCCCGGGGCTTCCCGTGTGCAAGACATCTTGAA 5021
Qy 4681 TTTTGGAGGAGATTTCACCGGCTCTACACACATAGATGCCACTTCTCTTCCCAACA 4740
Db 5022 TTTTGGAGGAGGCTCTTACCGGCTCTACATATAGATGCCACTTCTTATCCCAACA 5081
Qy 4741 AAGCAATGGGGGAAAAATTTGGCATACTTAACAGCTTACAGGCTACAGTGTGGCTTAG 4800
Db 5082 AAGCAGATGGGGAACCTTCTTACCTGTAGAGTACCAAGCACCGTGTGGCTTAGG 5141
Qy 4801 GCCAAAGCGCCCGCCGCTCTGAGGACGTCATGTGAGTGTGTGACTCGACTCAAGCC 4860
Db 5142 GCTCAGGCGCTTCCCGCATCGTGGGACCAAGTGTGAAAGTGTGTGATCCGCTTAAACC 5201
Qy 4861 AACTCTGTGGGCCCAACCTCTCTGTACCGCTTGGGCTCTGTATCAACAGAGTCAAC 4920
Db 5202 ACCCTCCATGGGCAACACCCCTGCTATACAGACTGGGCGCTGTTCAGAAATGAAAGTACC 5261
Qy 4921 CTCACACATCCCGTACGAATATACATGGCCACTGACATGCAAGCGACCTTGAAGTATG 4980
Db 5262 CTGACGACCCCAATACCAAAATACATCAATGACATGTCGCGCGACCTGGAAGTGTGTC 5321
Qy 4981 ACCAGCATGGGTCTTGGSCAGGGGAGTCTTGGCGCGCTGCGCGCTATTTGGCTGGCG 5040
Db 5322 ACGAGCACTGGGTCTGTGGCGCGCTCTGCGTGTCTGCGCGCTATTTGGCTGTGCA 5381
Qy 5041 ACCGGGTGTGTGATCATGCGCGCTTGGACATTTAACAGCGAGCGGTGCTGGCGCG 5100
Db 5382 ACAGGCTGCTGTGATAGTGGGCGAGATCGTCTTGTCCGGGAAGCCGGCAATTATACT 5441
Qy 5101 GACAAGAGGTCTCTATAGAGCTTTTGAAGATGAGAGGAATGTGCTCTAGAGGCGGCT 5160
Db 5442 GACAGGGAGTCTCTTACAGAGTTCATAGAGAGGAAGTGTCTCAAGCACTTACCG 5501
Qy 5161 CTCATTGAAGAGGCGACGAGTACCGAGATGCTGAAGTCCAGATTCAGAGCTTATTTG 5220
Db 5502 TACATCGACAGAGGATGATGCTGCTGAGCACTTCAAGCAGAGAGGCGCTGCGCTCTG 5561
Qy 5221 CAGCAGGCTTCCAAACAGCTCAGACATACCAACCACTGTGACAGGCTTCAATGGCCAG 5280
Db 5562 CAGACCGGCTCCCGCATGACAGAGGTTATCACCCCTGCTGTCCAGACCACTGCGAGAA 5621
Qy 5281 GTAGAACAATTTCTGGGCGAAACATATGTGAACCTTATACCGGATTCGATTACTCTGCA 5340
Db 5622 CTGAGGCTTTTGGGCGAAGACATGTGAAATTTATCATGAGGATTAACAATACTTGGCG 5681
Qy 5341 GGACATATCAACACTGCCAGGGAACCTGCAAGTCTTCATGATGGCTTCAAGTGGCGCC 5400
Db 5682 GGCTGTCAACGCTGCCGTGTGTAAACCCCGCATTTGCTTATGTATGCTTTTACAGTGC 5741
Qy 5401 CTCACAGTCCGCTGTCAACAGACCACTATCTTCTCAACATTTTGGGGGCGCTGGCTA 5460

Db 5742 GTACACAGCCCACTTAACCACTGGCCAAACCTCTCTTCAACATATTTGGGGGCTGGTG 5801
Qy 5461 GCATCCCAATTTGACACACCCCGGGGCGCACTGGCTTGTGTACAGTGGCTAGTGGGA 5520
Db 5802 GTGCGCAAGTCCGCGCCCGGCTGCTACTGCTCTTGTGTGTGTGCTGCTGAGCTGGC 5861
Qy 5521 GCTGCCGTAGGCGAGTATAGCTTATAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 5580
Db 5862 GCGGCATCGGACAGCGTTGAGTGTGGGAAGTGTCTGTGGAGATTTCTTGCAAGGATATGG 5921
Qy 5581 GCGGGCATTTGGGGGCTCTCGTGCATTTCAAGATATGTCTGGCGAAGAACCTTCCATG 5640
Db 5922 GCGGGGTGTGGCGGAGCTCTTGTAGCATTTCAAGATATGAGCGGTAGGTCCCTCCAG 5981
Qy 5641 GAGATGTGTCAACTTGTGCTGAGTGTGAAATTTGTCTCCGGGATGCTTGTAGTGGAGTGC 5700
Db 5982 GAGGACTGTGTCAATGTGCTGCGCCGCACTTCTGCGCTGTGAGGCTTTGTATGTGATGTG 6041
Qy 5701 ATCTGGCGGCAATTTCTGGCGCGACAGTGGGACCGGGGGAAGCGCGCTCAATGAGATG 5760
Db 6042 GTCTGGCAGCAATTAATGCGCGCGGACGTTGGCCGGGCGAGGGGCAATGTGCAATGGATG 6101
Qy 5761 AATAGACTATTTGCTTGTGTCTTCAAGAGAAATACAGTGTGCGCCCACTACCTGAGAG 5820
Db 6102 AACCGCTATATACCTTGTGCTTCCGGGGGAACATGTTTCCCCACGACATACGTGCGG 6161
Qy 5821 GAGTGGATGCGGTGTGAGCGGTGTGACCAACTGATCTTGGCTGCTTCAATTAACAGCTG 5880
Db 6162 GAGAGGATGTAGCGCCCGCGCTGATCTGCACTACAGACCTTCACTGTAAACCAAGCTTC 6221
Qy 5881 CTCAGAAAGCTCCCAACTGTGATTAATCTAGAGACTGCCCCATCCATGCGGCGCTGTGG 5940
Db 6222 CTGAGGAGACTGTACATAGTGTATAGCTGTGAGTGTATACCACTCACTGCTCGGTTCTGG 6281
Qy 5941 CTCGCGATGTGTGGGACTGTGGTTTGACCATCTTAACAGACTTTTAAATTTGGCTGACC 6000
Db 6282 CTAAAGGACATCTGTGGACTGTGATGTGAGTGTCTAGGCACTTTTAAAGCTGTGGGAA 6341
Qy 6001 TCCAAATTAATTTCCCAAGATATGCCGCGCTCCCTTGTCTCTGTCAAAAAGGGGTCAAG 6060
Db 6342 GCCAAGCTATGTCCACAACTGCTTGGGATTTCCCTTGTGTCTGCGACGCGCGGGTATAG 6401
Qy 6061 GCGGTGTGGCGCGGCACTGTGCATGACCAACGAGTGTCTTGGCGCGCAATATCTCT 6120
Db 6402 GGGGTGTGGGAGAGACCGCATTTATGACACTGTGCTGCACTGTGAGTGTGATCATCT 6461
Qy 6121 GGCATATGTCCGCTTGGGCTTCATGAGAAATTAACGGGCTTAAGCTGTGATTAATCTGG 6180
Db 6462 GGAATGTCAAAAACGGGACGATGAGGATGTGTGCTTAAGACCTGTGAGGAACATGTGG 6521
Qy 6181 CAGGGGACCTTCTCTATCAATTTGTTACACGGAAGGCGAGTGTGCGGAAACCGGCGCA 6240
Db 6522 AGTGGAGCTTCCCATTTAAGCTTACACGAGGCGCCCTGTATCTCCCTTCTGTGCGCG 6581
Qy 6241 AACTTTAAGTGTGCATCTGTGAGGATGTGCGGCTCAGAGTACGAGGATGACGAGCAC 6300
Db 6582 AACTATTAAGTGTGCGGTGTGAGGATGTGTGAGAGGAATATGATGAGATTAAGCGGGTG 6641
Qy 6301 GGGTCATACCATACATACAGAGACTCAACACTGATTAATTTGAAAGTCCCTTCCCACTA 6360
Db 6642 GGGGACTTTCACATGATGTGGGTATGACTACTAGAACATTTTAAATGCGCGTCCAGATC 6701
Qy 6361 CCGTCTCCGAGTCTTTTCTGGGTGAGCGAGTGTGAGATTCATAGGTTTGGCCCCACA 6420
Db 6702 CAATGCCCGAATTTTTCACAGAAATTTGAGAGGGGTGTGCTTACAGGTTTGGCCCCCT 6761
Qy 6421 CCGAAGCGTTTTTCGGGATGAGTCTCGTTCTGTGTGGGCTTAATTTGATTTGTCTG 6480
Db 6762 TGCAAGCTTGTGCGGGGAGAGATATCAATTAAGATAGAGCTCCACGAGTACCGGGTG 6821
Qy 6481 GGGTCCAGCTTCTTGTGCACTGTGAACCGGACACAGATATTTGATGTCAATGTCTAA 6540

Db 6822 GGGTCGAATTACTTTCGAGCCGGAACCGGACGTAGCCGTTGTTGACGTCCAGTCACT 6881
Qy 6541 GATTCATCTCATATCAGCGCGGAGACTGTCAGCGCGGCTTTAGCGCGGGGTCAACCCCA 6600
Db 6882 GATCCCTCCCATATTAAGAGAGAGCGCGCGGAGAAAGTTGGGAGAGGGGTCAACCCCT 6941
Qy 6601 TCCGAGGCAAGCTCTCGCGAGCGCAGCTATTCGAGCACTATGCTGCGAGCACTTGAC 6660
Db 6942 TCTATGGCCAGCTCTCGCTAGCCAGCTGTCCGCTCATCTCTCAAGGCACTTGAC 7001
Qy 6661 ACCCAGCGCAAAAGCTATGATGAGACATGGTGAATCTTAACCT- - - - -GTTCC 6708
Db 7002 GCCAACCATGATCTCCCTGACGCGGAGCTCATAGAGCTTAACCTCTGAGGAGCAAGAG 7061
Qy 6709 ATGGGGGGCGATGATGACTCGGATGAGTCTGGGTCCAAAGTGGTCTTTCGAGATCTCTC 6768
Db 7062 ATGGGCGGCAACATCACAGGGTTGAGTCAAGAGAAAGTGGATTCGAGATCTCTTC 7121
Qy 6769 GACCCAATGCTCGAAGAAAGAGCGCACTTGAGCCTTCGATACCATCAGAAATACATGCTC 6828
Db 7122 GATCCGTTTGGGAGAGAGAGATGAGCGGGAGGTCTCCGTAAGTGAAGAAATTCGCG 7181
Qy 6829 CCCAAGAAAGGTTCCCAAGCACTTTAAGCGCTGGGCAAGCGCTGATTAACAACCAAG 6888
Db 7182 AAGTCTGGAGATTCCGCGGGCCCTGCGTGGGCGGCGGAGCTAACACCCCG 7241
Qy 6889 CTGTGGAATCTGAGAAAGAGCCAGATTACCAACCGGCACTGTGGGGGTGTGCTCTC 6948
Db 7242 CTAGTAAAGACGTGAGAAAGAGCTGATCAAGAACCACTGTGTCATGCTGCGCTCA 7301
Qy 6949 CTTCTCTCTAGAAAACCCGAGCGCTCCCAAGAGGCGCGGAGAGTGGGCTTAAGT 7008
Db 7302 CCACTCTCAAGGTCCCTCTGTCGCTCCGCTCGGAAAAAGGTAGGTGTCTCAAC 7361
Qy 7009 GAGGACTCCATAGAGAGATGCTTCAACAGCTGGGCTTAAGTCTTGGCCAGCCCGC 7068
Db 7362 GAATCAACCTATCTACTGCTTGGCGAGCTTGSCACAAAGTTTGGCAGCTCTCA 7421
Qy 7069 CCAAGCGGCAATTCAGGCTTTCCAGCGGGGCGGGCGCTGCCAATTCGGGAGTCAAGAC 7128
Db 7422 ACTTCGGGCAATTAAGGGCGCAATTAAGCAAACTCTCTGAGCCCGC- - - - - 7469
Qy 7129 CCTCTGATGATGAGTGGCCCTTTGAGAGAGAGTTCCATCTCTTCATGCCCCCTCGAG 7188
Db 7470 CTTCTGGCTGCCCCCGAGCTCCGAGCTGATCTTATCTTCATGCCCCCTTGAG 7529
Qy 7189 GGGGAGCTTGAAGATCCAGACTTGAAGCTGAGCAGGTAGAGCCCCCCCCCAG 7248
Db 7530 GGGGAGCTGGGGATCCGA- - - - - 7549
Qy 7249 GGGGGGGTGGCAGCTCCCGGCTGGAGCTCGGGGTCTGTCTACTTGTCCAGAGAGAC 7308
Db 7550 - - - - -TCTCAGCGACGGGTCAATGTGTCAGGTCACTGAGGGCGGACAGC 7595
Qy 7309 GATCCGTCGTGTGCTGCTCATGTCACTCCGTCGAGCGGGGCTCAATTAACCTTGT 7368
Db 7596 GAAAGATGTCTGTGTCTCTCAATGTCTTATCTTCTGAGCAGCGCATCTGCAACCCGTC 7655
Qy 7369 AGTCCGGAAGAGAGATTACCGATTAAACCCCTTGAACAATCCCTGTTGCGATATCAC 7428
Db 7656 GTGCGGAGAGAAACAAAATGCGCCATCAACGCACTGAGCAATCGTTGCTACGCAATCAC 7715
Qy 7429 AACAGGTGTACTGTACCAACAAGAGCGCTCACTAAGGGCTAAAGAGTAACTTTT 7488
Db 7716 AATCTGTGTATTCACACACTTCAACGAGTGTGCTGCAAGGAGCAAGAAAGTCAATTT 7775
Qy 7489 GATAGATGCAAGTCTGCACTCCTACTACGACTCAGTCTTAAAGACATTTAAGTACGCG 7548
Db 7776 GACAGACTGCAAGTCTGCAAGCATTACAGGACGTGCTCAAGAGGTGTCAAGCAGCG 7835
Qy 7549 GCTCCCAAGGTCAACCGGAGCTCTCAACATGAGAGAGGCTTGGCAGTTAACCCCAAC 7608
Db 7836 GCGTCAAAAGTGAAGGCTTAATCTGTCTATCCGTAGAGAGAGCTTGCAACCGCCCCCA 7895

Qy 7609 CATTCGCAAGATCTAATAATGAGTTTGGGCTTAAGAGGTCCGCACTTGTCCGGAGG 7668
Db 7896 CATTCAAGCCAAATCAAGTTTGGCTATGAGGCAAAAGACTCCGTGTCATGCCAGAAAG 7955
Qy 7669 GCGGTTAACCAATCAAGTCCGTGTGAAGAGACTCTCTGAGAGACTTCAGAAACCAATT 7728
Db 7956 GCGGTAGCCCAATCAACTCGTGTGAAGAGACTTCTGGAAGAGAGTGAACCAATTA 8015
Qy 7729 CCCAACAATTAATGCGCAAAATAGAGTGTCTGCTGGAACCCACCAAGGGGGGCAAG 7788
Db 8016 GACACTACCAATGCGCAAGAAAGAGTGTCTGCTTACGCTGAGAAAGGGGGGTCTG 8075
Qy 7789 AAGCAGCTGCGCTTAATCTGTTAACCCTGACCTCGGCTGAGGGGTCTGAGAAAGTGGCC 7848
Db 8076 AAGCAGCTGCTCATCTGTTTCCCGACTTGGCGGTGCGCTGTGCAAGAGATGGCT 8135
Qy 7849 CTTATGACATTAACAAAACTTCTCAGCGGAGTGAAGGGGCTTCTTAATGATTCAG 7908
Db 8136 CTGTAGACGTGTGTAGCAAGCTCCCGTGGCGGTATGGAAGCTCTAAGGATTCGA 8195
Qy 7909 TATTCCCCGCTCAGCGGGTGAAGTTCTTGAAGCATGGCGGAAAGAGACCTT 7968
Db 8196 TACTCAACAGAGAGCGGGTGAATCTCTGTCAAGCGTGAAGTCCAAAGAGCCCG 8255
Qy 7969 ATGGGTTTTGATATGATACCGGATGCTTGAATCAACGCTCACTGAGAGACATCAG 8028
Db 8256 ATGGGTTCTGTATGATACCGCTGTGTTGATCTCAACATCACTGAGACATTCCTG 8315
Qy 8029 ACTGAGAGTCCATATATCGGGCTGCTCTTCCGAGAGAGGCCCACTGCTATAC 8088
Db 8316 ACGAGAGAGGCATTTTACCAATGTGTGACCTGAGACCCCAACCCGCGTGGCATAG 8375
Qy 8089 TCGCTAACTGAGAGACTTACGTGAGAGGCTTATGTTCAACAGCAAGGCCAACTGCT 8148
Db 8376 TCCCTCACTGAGAGGCTTATGTTGGGGGCTCTTACCAATTCAGGGGGAATCTGC 8345
Qy 8149 GGGTACAGGCGTGGCGCGCAGCGGGGTGCTCAACATAGCAGTGGGAAACATCACA 8208
Db 8436 GGGTACCGAGGTGCGCGGAGCGGAGCTGATCAACATAGCTGTGTACACCTTACT 8495
Qy 8209 TGCTAGTGAAGCTTATGAGCGCTTGAAGCTGCAAGGAGTAATCGCGCCCAATGCTG 8268
Db 8496 TGCTAATCAAGGCCCGGGAGCTGTCAAGCGGAGGGTCCAGAGCTGCAACATGCTC 8555
Qy 8269 GTATGCGGCGATGCTTGTGTCTCATCTCAAGAAAGCCAGGAGACCGAGAGCGG 8328
Db 8556 GTGTGTGCGAGCACTTAATGCTTATCTGAAAGTGGGGGTCCAGAGAGAGCGCGCG 8615
Qy 8329 AACCTGAGGCTTCAACGAGGCTATGACCAAGGTATTTGCGCCCTCTGCTGAGACCCCGC 8388
Db 8616 AGCTTGAAGCTTCAACGAGGCTATGACCAAGGTACTCGCCCCCGGGAGACCCCA 8675
Qy 8389 AGACCGGAGTATGATGAGAGCTGATTAACATCTTGTCTCTCAATGTGTGTGAGCGCTG 8448
Db 8676 CAACCAAGAAATGCACTTGAAGCTTATTAACATATGCTCTCCAAAGTGTAGTGGCCAC 8735
Qy 8449 GGGCCCAAGGCGCGCAGATTAATCTGAACAGAGACCTTAACACTCAATGCGCGG 8508
Db 8736 GACGCGCTGGAAGAGGGGTCTACTACTTAAACCCGTGACCTTAACCCCTCGCGAGGA 8795
Qy 8509 GCTGCTGGGAAACATTTAGACATCCCGTGTCAATTTGATGCTGGGAAACATATCAG 8568
Db 8796 GCGCGTGGAGACACAGACACATCTCAAGTAAATCTGCGTGAAGCAATATCATG 8855
Qy 8569 TACGCCCCGACATATGAGGCTCGCATGTCTGATGACACACTTCTTTCATCTTCATG 8628
Db 8856 TTTGCCCCCAACTGTGGGAGAGATGATATGATGACCATTTCTTTAGCGTCTCATTA 8915
Qy 8629 GCTCAAGACAGCTGAGACCAAGACCTCAACTTTGAATGTAGAGAGCGGTGTACTCGGT 8688
Db 8916 GCGAGGATCAAGCTTAAAGAGGCTTAACTGTGATATCAAGGAGCTGTGCTCATTA 8975

QY 8689 AGTCCCTTGACCTCCAGCTAATATTGAAGTTACATGCGCTTTTCTCTG 8748
DB 8976 GAACCACTGATCTACCTCCATCATATCTCAAGACTGCGCTCAGCGCATTTTCACTC 9035
QY 8749 CACACATTCATCTCCCAAGAACTGACAGCGGTGCTTACGCGCTTCAAAAACTTGGGCGC 8808
DB 9036 CACAGTTACTCTCCAGGTGAATCAATAGGGGTGCGCGGATGCTTCAGAAAACTTGGGCTC 9095
QY 8809 CCACCCCTCAGAGCGTGAAGAGCGCGACGTCAGTCAGGAGCGTCCCTCATCTCCCGT 8868
DB 9096 CCGCCTTTCGAGCTTGAAGACACCGGCGCGGAGCGTCCCGCTAGGCTTCTCTCAGA 9155
QY 8869 GGGGGGAGCGCGCGCTTTGCGGTGATATCTTTCAATTGCGCGGTGAAGACCAAGCTC 8928
DB 9156 GGAGGAGCGGCTGCTAATATGTGCAAGTACCTTTCAACTGGGAGTAAGAAACAAGCTC 9215
QY 8929 AAATCTCATCTCATTTGCGGAAAGCGCGCTCTGATTTATTCAGCTGTTTACCGTGGC 8988
DB 9216 AAATCTCATCTCAATAGCGGCGCTGCGCGCTGAGCTTGTCCGTTGTCACGCGTGC 9275
QY 8989 GCGGCGGCGGCGACATTTATCAAGCGTGTGCGTGCACCGCGCTTATGCTCTT 9048
DB 9276 TACAGCGGCGGAGACATTATACAGCGTGTCTATGCGCGCGCGCTGCTTCTGGTTT 9335
QY 9049 GGCCTACTCTTCTTTTGTAGGGGTAGCGCTTTTCTTACTCCCGCTCG 9098
DB 9336 TGCTACTCTGCTGCTGAGGAGTAGGATCTTACTCTCCCAACCG 9385

RESULT 9

AAC86646
ID AAC86646 standard; DNA; 9611 BP.

AC AAC86646;

XX 02-APR-2001 (first entry)

DE Nucleotide sequence of chimeric Hepatitis C virus clone pH77 (p7)-CV-J6S.

XX HCV, HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine; ss.

OS Synthetic.

OS Hepatitis C virus.

XX Key Location/Qualifiers

FT CDS 342..9389

XX P-PSDB; AAB30731.

XX WO200075338-A2.

XX 14-DEC-2000.

XX 02-JUN-2000; 2000WO-US015446.

XX 04-JUN-1999; 99US-0137693P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Yanagi M, Bulk J, Emerson SU, Purcell RH;

XX WPI; 2001-061728/07.

XX P-PSDB; AAB30731.

XX Nucleic acid molecule encoding human hepatitis C virus of genotype 2a for

XX assays for identification of antiviral agents.

XX Diaclosure; Page 115-118; 167p; English.

XX AAC86645-48 represent chimeric cDNA clones of infectious Hepatitis C

XX virus (HCV). In each clone the C, E1 and E2 genes are derived from HCV

XX strain HC-J6CH genotype 2a. The p7 protein was derived from HCV strain HC

CC strain PCV-H77C. Such HCV sequences are capable of expressing the virus
CC when transfected into cells. The HCV protein is useful for assaying
CC candidate antiviral agents for activity against HCV. Antibodies specific
CC for HCV polypeptide are useful in prevention and treatment of diseases
CC caused by HCV in animals, in particular humans. The HCV polypeptides
CC serve as immunogens in the development of vaccines for preventing HCV in
CC mammals or as antigens in diagnostic assays for detecting the presence of
CC HCV in biological samples. The HCV polynucleotide is also useful for
CC identifying cell lines capable of supporting the replication of HCV in
CC vitro and to produce attenuated viral strains via passage in vitro or in
XX vivo

SQ Sequence 9611 BP; 1924 A; 2898 C; 2694 G; 2095 T; 0 U; 0 Other;

Query Match 58.0%; Score 5274.4; DB 5; Length 9611;

Best Local Similarity 74.1%; Pred. No. 0;

Matches 6751; Conservative 0; Mismatches 2281; Indels 78; Gaps 3;

QY 1 ATGAGCAAAATCTTAACCTCAAGAAAAAACCAGAAACCAACCGTGGCCACAA 60
DB 342 ATGAGCAAAATCTTAACCTCAAGAAAAAACCAGAAACCAACCGTGGCCACAA 401
QY 61 GAGTTAAGTTTCCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
DB 402 GAGTTAAGTTTCCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 461
QY 121 GCGCCAGAGTTGGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
DB 462 GCGCCAGAGTTGGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 521
QY 181 AGGCGCGAGCCATCCCTTAAGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
DB 522 AGGCGCGAGCCATCCCTTAAGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 581
QY 241 TACCCCTGCGCCCTTAAGCGGAAATGAGGAACTCGCGTGGCGAGGATCTTCTGTC 300
DB 582 TACCCCTGCGCCCTTAAGCGGAAATGAGGAACTCGCGTGGCGAGATGCTCTGTC 641
QY 301 CGAGTTCCCGTCCCTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 642 CGAGTTCCCGTCCCTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 701
QY 361 AAGTCATGATACCTTAAGCGGAAATGAGGAACTCGCGTGGCGAGGATCTTCTGTC 420
DB 702 AAGTCATGATACCTTAAGCGGAAATGAGGAACTCGCGTGGCGAGGATCTTCTGTC 761
QY 421 GCGGCG 480
DB 762 GCGGCG 821
QY 481 GGGGTTAATTTTGAACAGGAACTTAACCGGTTGCTCCCTTTCTTCTTCTTCTG 540
DB 822 GGGGTTAATTTTGAACAGGAACTTAACCGGTTGCTCCCTTTCTTCTTCTTCTG 881
QY 541 CTGCTGCTGATCAACCAACCGGTTGCTCCCTTTCTTCTTCTTCTG 600
DB 882 CTGCTGCTGATCAACCAACCGGTTGCTCCCTTTCTTCTTCTTCTG 941
QY 601 TACATGATGATCAACCAACCGGTTGCTCCCTTTCTTCTTCTTCTG 660
DB 942 TACATGATGATCAACCAACCGGTTGCTCCCTTTCTTCTTCTTCTG 1001
QY 661 CTCACAGTCCCGGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
DB 1002 CTCACAGTCCCGGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1061
QY 721 CCGGTTCAACCGAATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
DB 1062 CCGGTTCAACCGAATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1121
QY 781 CACATGACATGTTGATGATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
DB 1122 CACATGACATGTTGATGATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1181

QY 841 TCGGTGGGGGATGCTCGAGCCCAATGTTCAATTGTCGCGGACCACTGGTTT 900
DB 1182 TCGGTGGGGGATGCTCGAGCCCAATGTTCAATTGTCGCGGACCACTGGTTT 1241
QY 901 GTCCAGATGCAATGCTCCATCTACCTTGTAACATCACTGACACCGCATGGAC 960
DB 1242 GTCCAGATGCAATGCTCCATCTACCTTGTAACATCACTGACACCGCATGGAC 1301
QY 961 GACATGATGATGATGCTGGGCGGACCTGATGATCTTGAGGTACGGATGCTGTC 1020
DB 1302 GACATGATGATGATGCTGGGCGGACCTGATGATCTTGAGGTACGGATGCTGTC 1361
QY 1021 CCCAGGTCAATTATAGATCAATTAAGCGGGCTCATTTGGGGCGTATGTTCCGTTGGCC 1080
DB 1362 CCCAGGTCAATTATAGATCAATTAAGCGGGCTCATTTGGGGCGTATGTTCCGTTGGCC 1421
QY 1081 TACTTCTATGACGAGGAGCGTGGCGAAGTCGTTGTCATCTTCTGTTGGCCGCGG 1140
DB 1422 TACTTCTATGACGAGGAGCGTGGCGAAGTCGTTGTCATCTTCTGTTGGCCGCGG 1481
QY 1141 GTGACCGCGGACCCATACCTGTTGGGGGTTCTGCGCGACACCGGGCGCTTACC 1200
DB 1482 GTGACCGCGGACCCATACCTGTTGGGGGTTCTGCGCGACACCGGGCGCTTACC 1541
QY 1201 AGCTTATTTGACATGAGGCGCCGACAGAAATCCAGCTCGTTACACCAATGGACGTGG 1260
DB 1542 AGCTTATTTGACATGAGGCGCCGACAGAAATCCAGCTCGTTACACCAATGGACGTGG 1601
QY 1261 CACATCAACCGCACCGGCTTGAACTGCAATGACTCTTTCGACACCGGCTTTATCGCTCT 1320
DB 1602 CACATCAACCGCACCGGCTTGAACTGCAATGACTCTTTCGACACCGGCTTTATCGCTCT 1661
QY 1321 CTGTTCTACACCCACAGCTTCAACTCGTACAGATGTCGCCAGACGATGTCGCTGCGCG 1380
DB 1662 CTGTTCTACACCCACAGCTTCAACTCGTACAGATGTCGCCAGACGATGTCGCTGCGCG 1721
QY 1381 AGTATCGAGGCTTTCCGGGTGGAGTGGGGGCGCTTGCAATATGAGATTAATGTCACCAAT 1440
DB 1722 AGTATCGAGGCTTTCCGGGTGGAGTGGGGGCGCTTGCAATATGAGATTAATGTCACCAAT 1781
QY 1441 CCAGAGATATGAGACCTTATGCTGACGACTACCCACCAAGGAGTGGCGTGTCTCC 1500
DB 1782 CCAGAGATATGAGACCTTATGCTGACGACTACCCACCAAGGAGTGGCGTGTCTCC 1841
QY 1501 GCGAAGACTGTGTGGGCCAGTGTACTGTTTCAACCCCAAGCCAGCTGTGTGGGACG 1560
DB 1842 GCGAAGACTGTGTGGGCCAGTGTACTGTTTCAACCCCAAGCCAGCTGTGTGGGACG 1901
QY 1561 ACCGACAGGCTTGGAGGCGCCACTTACACGATGGGGGAGATGAGACAGATGTCCTCTA 1620
DB 1902 ACCGACAGGCTTGGAGGCGCCACTTACACGATGGGGGAGATGAGACAGATGTCCTCTA 1961
QY 1621 TTGAACAGCACTGACACCGCTGGGGTCAATGTTTCGCTGACGATGATGAATCTTCT 1680
DB 1962 TTGAACAGCACTGACACCGCTGGGGTCAATGTTTCGCTGACGATGATGAATCTTCT 2021
QY 1681 GGGTACACCAAGACTTGGGGGACACACCTGCGCTTCTAGAGCTGACTTCAACGCGACG 1740
DB 2022 GGGTACACCAAGACTTGGGGGACACACCTGCGCTTCTAGAGCTGACTTCAACGCGACG 2081
QY 1741 ACCGACAGGCTTGGGCGCCACGAGCTGTTTATGAGAGATCTCTATACCACTTACCTCAA 1800
DB 2082 ACCGACAGGCTTGGGCGCCACGAGCTGTTTATGAGAGATCTCTATACCACTTACCTCAA 2141
QY 1801 TGGGGCTCTGGGCGCTTGGCTCACGCGCAAGGTGCTGATGCACTAACGCTCTGG 1860
DB 2142 TGGGGCTCTGGGCGCTTGGCTCACGCGCAAGGTGCTGATGCACTAACGCTCTGG 2201
QY 1861 CATTAACCCCTGACAGTATCTATACATCTTCAAAATAGATGATGTTGGGAGGGGTT 1920
DB 2202 CATTAACCCCTGACAGTATCTATACATCTTCAAAATAGATGATGTTGGGAGGGGTT 2261

QY 1921 GAGACAGGCTCAAGGCTGATGCAATTTTCACTGATGGGAGTGTGCAACTTGAGAGAC 1980
DB 2262 GAGACAGGCTCAAGGCTGATGCAATTTTCACTGATGGGAGTGTGCAACTTGAGAGAC 2321
QY 1981 AGAGACAGAAATGCACTGTCTCTTGTGTCACCTCACACGGAATGGGCAATTTTAACT 2040
DB 2322 AGAGACAGAAATGCACTGTCTCTTGTGTCACCTCACACGGAATGGGCAATTTTAACT 2381
QY 2041 TGCCTTAACTCGGACCGCGCGCTGTGCACTGTGCTTTTCCACCTTCCACCAAAATATC 2100
DB 2382 TGCCTTAACTCGGACCGCGCGCTGTGCACTGTGCTTTTCCACCTTCCACCAAAATATC 2441
QY 2101 GTGAGCGTACAAATTCATATGATGAGCTATCACCTGACCTCAAAATATGATGTCGATG 2160
DB 2442 GTGAGCGTACAAATTCATATGATGAGCTATCACCTGACCTCAAAATATGATGTCGATG 2501
QY 2161 GAGTGGATTAATCTTATTTCTGCTTATAGCGGACGCGAGGTTTGGCGCTGTATGG 2220
DB 2502 GAGTGGATTAATCTTATTTCTGCTTATAGCGGACGCGAGGTTTGGCGCTGTATGG 2561
QY 2221 ATGCTCATCTTGTGGGCGCGGACGAGACGACTAGAGAACTGTGCTATCTTGACAGCT 2280
DB 2562 ATGCTCATCTTGTGGGCGCGGACGAGACGACTTGGAGAACTGTGATATCTCAATGCA 2621
QY 2281 GCGAGCGAGCTAGCTGCAATGGCTTCTTATATTTTGTCACTTTTTCGTTGGCGCTTGG 2340
DB 2622 GCAATCCCTGGCGGAGACGACGAGCTTGTGTCCTTCTGCTGTGTTCTTGGCTTGG 2681
QY 2341 TACATCAAGGCTGGGATGATCCCTTATAGCTAATCTTATCCCTCACTGGCGTGTGCTT 2400
DB 2682 TATCTGAAGGATGATGAGGATGCGGAGCGGCTTACGCTTACGGAATGAGGCTTCTC 2741
QY 2401 AGCTTACTGCTCTAGATTTGCCCCAACAGCTTATGCTATGACCATGTCGATGCG 2460
DB 2742 CTCTCTCTCTGCTGCGGTTGCTTCAAGCGGCAATGCACTGACACGAGATGCGCGG 2801
QY 2461 CAGATAGAGCGGCTGCTGCTGATATGATCACTCTTATCTCTCAACCCCGGATTAAG 2520
DB 2802 TCGTGTGGCGGCGTGTGTTCTGTCGGGTTAATGGCGCTGACTCTTGTGCGCATTTACAG 2861
QY 2521 ACCCTTCTAGCGCGGTTTGTGATGTGTGCTATCTTCTGACCCCTGGGGAAGCTATG 2580
DB 2862 CGCTATATACGCTGTGATGTGTGCTTCACTATTTTCTGACAGAGATGAAGAGCGCAA 2921
QY 2581 GTCCAGAGATGGGACCACTTATGACAGTGTGCGGTTGCGCTGATGACATATGGGCC 2640
DB 2922 CTGCAAGTGTGGGTTTCCCCCTCAACGTCGCGGGGGGGCGCATGCTGCTTATCTC 2981
QY 2641 GTGCGCATATTTACCCAGGTGTGTGTTGACATTAACCAAGTGGCTTGTGGGCGTCT 2700
DB 2982 ATGTGTGTATACCCGACCTGTGTATTTGACATACCAAACTACTCTTGGCGATCTTC 3041
QY 2701 GGGCGTGTCTTACCTCTTAAAGGTGCTTGAACGCGGTGCGTACTTGTCAAGGCTCAC 2760
DB 3042 GGAACCTTGTGATTTCTTCAAGCGACTTGTCTTAAAGTCCCTTACTTGTGCGCGTTCAA 3101
QY 2761 GCTCTACTAGATGTGACACATGACAGGACTCTGCGGGGGGAGATGCTGCAAGT 2820
DB 3102 GGGCTTCTCCGATGTGCGGCTGAGCGCGGAATATGCGGAAGTCAATTAACGCAAAATG 3161
QY 2821 GGGCTACTAGCTTGGAGAGTGAATGCACTTATGATGACACCTCACCTATG 2880
DB 3162 GGCATATCAAGTTAGGGGCGCTTATCTGGCACTATGTGTATTAACATCTCACCTCTT 3221
QY 2881 TCGGATTTGGCTGTAGTGGCTTGGCGGACCTTGGCGGTGCGCTTGAAGCTTATCTTCT 2940
DB 3222 CGAGACTGGGCGGACAAAGGCTGCGAGATCTGGCGGTGTGGAACAGATGCTTCT 3281
QY 2941 AGTCCGATGGAAGAAATGATGCTGTTGTTGGGAGCGGACACGCTGTTGGGACAT 3000
DB 3282 TCCGATGGAAGCAAGCTCATACGATGGGGGAGATACGCGCGGTGCGATGAC 3341
QY 3001 TTACAGGACTTCCCGTGTCCGCGCACTTGTGCGGAGGTCTCTTGGCCGACGTGAT 3060

Db 3342 ATCAACGGCTTGCCTGCTCTGCGCCGTAAGGGCCAGAGATATGCTTGGGCCGAGCCGAC 3401
Qy 3061 GGCTAATCCTCCAAAGGGGTGAGTCTTCTGCCCCCAATCACTGCTTAAGCCAGACAGACA 3120
Db 3402 GGAATGCTCTCAAGGGGTGAGGTTGCTGGGCCCATCAAGGCTACGCCACAGCAAGC 3461
Qy 3121 CGTGGCCCTTTTGGGCAACATAGTGGTGAAGATGAAGGGGCGGACAAAGACAGAGCT 3180
Db 3462 AGAGGCCCTCTAGGGTGTATATCAACGAGCTGAGCTGGCCGGGACAAAGAACAGTGGAG 3521
Qy 3181 GGGGAAATTTGAGTCTCTGCTCAAGTCACTAGTCTTCTCTGGAACATCATCTCGGGG 3240
Db 3522 GGTAGGCTCAAGATCTGTCAACTGTCAACCAACCTTTCCTGGAACGTGTATTAATGGG 3581
Qy 3241 GTTTTGTGAGCTGTCTACATGAGAGCTGGCAACAAAGTGGCCGGCTCAAGGGTCCG 3300
Db 3582 GTATGCTGAGCTGTCTACACGGGGCCGGAAACAGAGCAATCGATCAACCAAGGCTCT 3641
Qy 3301 GTCAAGCAGATGTACTCCAGTGTGAGGGGGACTTTAGTGGGTGGCCCAAGCCCTGGG 3360
Db 3642 GTCAATCCAGATGTATACCAATGTGGACCAAGACCTTGTGGGCTGGCCGCTCTCAAGT 3701
Qy 3361 ACTAATCTTTGAGCCGTGACCGTGTGAAGCGGTGACCTGTACCTGTGTATAGCGGAAAC 3420
Db 3702 TCCGCTCATTTGACACCTGTACCTGCGGCTCTCGGACCTTTACCTGTGTACAGAGGAC 3761
Qy 3421 GCTGATGTATCCCGGCTCGAAGAGCGCGGGGACAAACGGGGAGGCTACTCTCCCGAGA 3480
Db 3762 GCCATGTATCTCCGTGTGCGCGGAGAGGTATAGCAGGGGTAGCTGTCTTCCGCCCGG 3821
Qy 3481 CCTCTTTCACCTTGAAGGGGTCTCTAGAGAGCCCGGTGTATGCCCCAGGGAGCACGCT 3540
Db 3822 CCAATTTCTCTATTGAAGGCTCTCTCGGGGGGTCTCGTGTGTGCCCCGCGGACACGCC 3881
Qy 3541 GTCCGAGCTTCCCGGAGCTGTGTGTCTGTGGGGCCGTGTAGTCCATAGATTTTATC 3600
Db 3882 GTGGGCTTATTCAAGGGCCGCGGTGTGACCCGTGAGTGTCTAAAGCGTGTGACTTATC 3941
Qy 3601 CCGGTGAGAGACTGACATGTCACGCGGTCCGCCACTTTAGTGAACAAGACAGACACA 3660
Db 3942 CTTGTGAGAGAACTAGAGGACAAACATAGAGATCCCGCGTTCACAGACAACTCTCTCCA 4001
Qy 3661 CCTGTGTGCCCCAGACCTATCAAGTGGGTACTTTGATGCTTGAAGCTGGAGTGAAG 3720
Db 4002 CCAGAGTGGCCCCAGAGCTTCCAGGTGGCCCACTGCTATGCCACCGGAGCGGTAAAG 4061
Qy 3721 AGCCAAAGTTTCTGTGAGCATATGCTGTCAAGGGGTAAAGTCTAAGTCTTAATGCC 3780
Db 4062 AGACAAAGTCTCCGGCTGCGTACGCAAGCCAGGGCTAACAGTGTGTCTCAACCCC 4121
Qy 3781 TCAGTGCTGCAACCTGGGGTTTGGGGCTGTACTTTGAAGGACATGATCATATCCC 3840
Db 4122 TCTGTGTGTCGAACCTGGGCTTTGTGTGTCTTCAATGTCAAGGCCCATGGGGTGTATCT 4181
Qy 3841 AACATTAAGACTGAGTCAAGACTGTGACGACCGGGGCGCCATCACTACTCTCAATAT 3900
Db 4182 AATATCAAGAACCGGGGTGAGAAACAATTAACCACTGGACGCCCATCACTATCCACCTAC 4241
Qy 3901 GGCAAATTTCTGCGCGGATGGGGGCTGTGCGGGCGGCTTACAGCATATCATATGTAT 3960
Db 4242 GGCAGTTTCTTGGCTGCGAGCGGGGTGTCAAGAGGTGTATGACATTAATATTTGTGAC 4301
Qy 3961 GAATGCTAGCCGTGAGCTTATACACATCTTGGCATCGGAAAGCTCTTGATCAAGCA 4020
Db 4302 GAGTGCACCTCCACGGAATCCACATCTTTGGCATCGGACCTGTCTTGACCAAGCA 4361
Qy 4021 GAGACAGCTGAGTCAACTATCTGTGTGTCTACAGCTACGCCCTTGGGTCAATGACA 4080
Db 4362 GAGACTGCGGGGCGAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4421
Qy 4081 ACCCCCAACCCCAATATAGAGAGGTGCGCTTGGGACAGAGGGGAAATCCCTTCTAT 4140

Db 4422 GTGTCCCATCTTAACATGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4481
Qy 4141 GGGAGGGGCAATTTCCCTGTCTTATCAATGAAGGAGAAAGACATGTATCTTGTCCATTTCA 4200
Db 4482 GGCAGAGCTATCCCTCCATGAGTATCAAGGGGGAAGACATCTATCTTGTGCCATCTCA 4541
Qy 4201 AAGAAAAAGTGTACGAGCTTCGCGCGGCCCTTTCGGGGTATGGGCTTGAATCTCATGTGCA 4260
Db 4542 AAGAAAGATGTGACAGAGCTCGCGGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4601
Qy 4261 TACTACAGAGGTGTGAGCGTCTCCGTAATACCAACTCAAGGGAGCGTATGTGTGTGTGT 4320
Db 4602 TACTACCGAGGTGTGACGT 4661
Qy 4321 ACCGAGCCCTCATGACAGGGTATATCTGGGGAATTTGACTTCGTATGATCTGCAACGTA 4380
Db 4662 ACCGATGCTCTATGATCTGTGCTTTCACGGGACTTTCAGCTGTGTATGATCTGCAACAG 4721
Qy 4381 GCGGTCACTCAAGTTTATCTTCAAGTTTATGACCCCAATTCACCTAACACACAGATT 4440
Db 4722 TGTGTCACTCAGACAGTGTGATTTCAAGCTTGAACCTTACCTTTACATGAGACACAG 4781
Qy 4441 GTCCCTCAAGAGCTGTCTCAAGTACGACCGCGGGGTGTGACCGGGTGAAGGAGACG 4500
Db 4782 CTCCCAAGAGTGTGTCTTCAGAGCTCAACCGCGGGGAGAGCTGAGAGGGAGAGCA 4841
Qy 4501 GGCATTTATAGTATGTTTCACTGTGTGAGCGAGCTCAGAGATTTTGTATGATGTATG 4560
Db 4842 GGCATCTATAGATTTGTGTGACACCGGGGAGCGCCCTCCGGGATGTGTGACTGTGTGT 4901
Qy 4561 CTCTGTGAGTGTATGACAGCGAGGGGCGCATGTGTATGAGCTCACACATTCGAGACAC 4620
Db 4902 CTCTGTGAGTGTATGACAGCGAGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4961
Qy 4621 GTGAGGCTCAAGGGGCTATTTCAACAGCGCGGGTGTGTGTGTGTGTGTGTGTGTGT 4680
Db 4962 GTTAGGCTACAGAGCTATGACATGACACCGGGGCTTCCGTGTGTGTGTGTGTGTGT 5021
Qy 4681 TTTTGGAGGCAAGTTTTCACCGGCTCAGACACATAGATGCCACTTCTTTCCAAACA 4740
Db 5022 TTTTGGAGGCGCTTTTACCGGCTCAGCTATATAGATGCCACTTTTTATCCAGACA 5081
Qy 4741 AAGCAATCGGGGAAAAATTTGCTATCTTAACAGCTTACAGGCTACAGTGTGTGTGT 4800
Db 5082 AAGCAGATGGGAGAACTTTCTTACCTGTATGCTTACCAAGCCACGTTGTGCGCTAAG 5141
Qy 4801 GCGAAAGCCCCCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4860
Db 5142 GCTCAAGCCCTTCCCATGT 5201
Qy 4861 ACACTGTGTGCCCCACACTCTCTGTATCCGCTTGGGCTGTGTATCCAAAGAGTCAAC 4920
Db 5202 ACCCTCATAGGGCAACACCCCTGTATATACAGCTGGGGCTGTTCAGATGAAAGTCAAC 5261
Qy 4921 CTACACATCTCCGTGACGAATATCATGCGCACTGTATGACGCGACCTTGAAGTCAAG 4980
Db 5262 CTGACGCAACCAATATCAAAATATCATATGACATGTATGTGCGCCCACTGTGAGGTCTG 5321
Qy 4981 ACCAGACATGGGTCTTGTGAGGGGAGTGTGTGGCGGCGGTGTGCGCGTATGTGCGG 5040
Db 5322 ACAGGACCTGGGTGTCTGTGGCGGCGTCTGGCGGTCTGTGCGCGGTATGTGCGTGTCA 5381
Qy 5041 ACCGGGTGTGTGTGATCATCTGCGCGCTTGCACATTAACAGGAGCGGTGTGTGCGCG 5100
Db 5382 ACAGGCTGTGTGTGTATGT 5441
Qy 5101 GACAGAGAGTCTCTATATGAGCTTTTGTATGATGAGAGAGATGTGTCTTATGGGCGCT 5160
Db 5442 GACAGGAGAGTCTCTATACGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5501
Qy 5161 CTGATTGAAAGAGGGGAGCGGATATGCGAGAGATGCTGAAAGTCCAAAGTCCAAAGTTAT 5220
Db 5502 TACATGAGCAAGAGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5561

QY	5221	CAGCAGCTTCCAAACAAAGCTCAAGACATACAAACCCTGCGAGGCTTCAATGCCCAAG	5280
Db	5562	CAGACCGGTCCCGCCATGCAAGAGTTATCAACCCCTGTGTCAGAACCAACTGGCAAGAA	5621
QY	5281	GTAGAACATTTCTGGGCCAAACAGATGGAATCTTACCGGCATCCATACCTCGCA	5340
Db	5622	CTCGAGGCTTTTGGGCGAAGACATGTGGAAATTTCACTAGTGGGATACATATCTTGGCG	5681
QY	5341	GGACTATTAACACTGCCCAGGGAACCTTGCAAGTAGCTTCATGATGGCGTTCAATGCCGC	5400
Db	5682	GGCTGTCAAGCTGCCCCTGGTAAACCCGCCCATTTGCTCATTAAGCTGCTTTTACAGCTGCC	5741
QY	5401	CTCAACAGTCCGCTGTCAACAAGCAACCACTATTCCTTCCAACTTTTGGGGGGCTGAGCTA	5460
Db	5742	GTCAACAGCCCACTAACCACTGCGCCAAACCTCTCTTCAACATATTTGGGGGGGTGGGTG	5801
QY	5461	GCATCCCAATTTGCAACCAACCCGCGGGGCCACTGGCTTCTGTTTCACTGAGGCTAGTGGGA	5520
Db	5802	GCTCCCAAGCTGCGCGCCCCCGGGTGCCTCTCTGCTTTTGGGGTGGCTGAGCTGAGCG	5861
QY	5521	GCTCCCGTAGGCAAGTATAGGCTTTAGTAAAGGTGCTAGTGGACATCTCGGCAAGGATATAGT	5580
Db	5862	GCCGCATCGGCAAGGTTTGGACTGGGGAGGATCTCTCGGACATTTCTTGCAAGGGATATAGC	5921
QY	5581	GCGGGCAATTTGGGGGGCTCTGTCGCAATTCAGATTCATGTCGCGAGAAACCCCTCATG	5640
Db	5922	GCGGGCGGTGGCGGAGGCTCTTGAGCATTTAAAGATCAGAGCGGTGAGGTCCCTCCACG	5981
QY	5641	GAGCATGTCTCAACTTGTCTGTGCGGAATTTGTGTCTCGGGTGCTTGGTATGTGGAGTGC	5700
Db	5982	GAGAACCTGTGTCAATCTGTCTGCCCCGCCATCTCTCGCTTGAGGCCCTTGTATGTGCGTGTG	6041
QY	5701	ATTCGCGCGGCATTTCTGCGCGCGCAACGCTGGGACCGGGGGAAAGCGCGCTCAATGTAGT	5760
Db	6042	GTCTGCGCAGCAATATCTGCGCCGGCAGTGTGACCCGGCCAGAGGGGCAATGCAATGATG	6101
QY	5761	AATAGACTCATTTGCTTTGTCTTCCAGAGGAATCAAGTCGCGCCCCCAACCCTACATGATGACG	5820
Db	6102	AACCGGCTAATAGCTTTGCTGCTCCCGGGGGAACATGTTTCCCAAGCACTTACGTGCGG	6161
QY	5821	GAGTCGATGTGGTGGCAGCGTGTGACCACTACTTGTGCTCCCTTACATTAACAAGCTG	5880
Db	6162	GAGAGCCATGTAGCGCGCCCGGTCACTGCACTCAAGCAACCTCACCTGTAAACCACTGC	6221
QY	5881	CTCAGAAGACTCCAAACTGTGATTAAGTAGGACTGCCCATCCGATCGCGCGGCTGTGG	5940
Db	6222	CTGAGGGGACTGCACTCAAGTGGATTAAGCTCGGAAGTATCACTCATGCTCCGGTTCTGTGG	6281
QY	5941	CTCCGCGATGTGGGACTGGGTTTGCAACATCTTAACAAGCTTTAAAAATTTGGCTAACC	6000
Db	6282	CTAAGGGACATCTGGGACTGGATGTGGAGGTGTGAGCGACTTTAAGACCTGTGCTAAA	6341
QY	6001	TCCAAATTAATCCCAAAAGATGCGCGGCTCCCTTTGTCTCTGTCAAAAAGGGGTACAAG	6060
Db	6342	GCCAAAGCTCATGCCACAACACTGCTGGGATTTCCCTTTGTCTCTGCAAGCGCGGGTATAGG	6401
QY	6061	GGCGTGTGGCGCGCACTGGCATATGACCAACAGGTGCTCTTTCGCGCGGCCAATATCTCT	6120
Db	6402	GGGGTCTGGCGAGAGAGCGGCAATTATGACACTCGCTGCACTGTGGAGCTGAGATCACT	6461
QY	6121	GGCAATGCCCGCTTGGGCTCCATAGAAATCAACGGGGCTTAAGACCTGCAATGAATATCTGG	6180
Db	6462	GGACATGTCAAAAACGGGACGATAGGATCTCGTCTCTTAGAGACTTGGAGGAACATGTGG	6521
QY	6181	CAGGGGACCTTTCCCTATCAATTTGTTTACACGAGAGGGACAGTGCCTGCGAACCCTCGCCCA	6240
Db	6522	AGTGGAGACTTTCCCATTAACGCTTACCAACGAGGGCCCTGTATCTCCCTTTCGTGGCGG	6581
QY	6241	AACTTTAAGGTGCGCATGTGAGAGGTGCGGCTTCAGAGTACGCGAGAGTATGACGACAC	6300
Db	6582	AACATATAGTTGCGCTGTGAGAGGTGTCTGCAAGAAATATCTGTGAGAGTATAGCCGGGTG	6641

OY	6301	GGGCTATACACATCAATAAACAGACCTCAACA	CTGATTAACCTTGAAGTCCCGCCAACTA	6360
Db	6642	GGGGACCTTCCACTACGATATCGGGTATGACTA	CTGACAACTTTAAATGCCCGTCCAGATC	6701
OY	6361	CCCTCTCCCAAGTCTTTTCTCGGTGTGACGAG	TGCAGATCCATAGTTTGGTCCCA	6420
Db	6702	CCATCGCCGGAATTTTTTCAACGAATTTGAA	AGGGGTGGCTACAGAGTTTGGGCCCT	6761
OY	6421	CCGAAGCCGTTTTTCCGGGATGAGTCTCGTT	CGTGGCTTAATTCATTTGTCTGC	6480
Db	6762	TGCAGGCCCTTGTCTGCGGGAGGAGTATCAT	TCAAGTAGAGTACCTCACAAGTACCCGGTG	6821
OY	6481	GGGGCCAGGCTTCCTTGGCCGCCCTGAACCG	CAACAGAGTATTTGATGTCAATNGCTACA	6540
Db	6822	GGGTGCAATTTACTCTTGCGAGCCCGAA	CCGAGTATGCGGTGTGACGTCAAGCTCACT	6881
OY	6541	GATCATCTATATATCAACGCGAGACATGAC	CGCGCGCTTTTGAACGCGGGGGTACCCCA	6600
Db	6882	GATCCCTTCATTAATACAGACGAGGGCG	CGCGAGAAAGTTTGGCGAGAGGCTACCCCT	6941
OY	6601	TCCGAGGCAAGCTTCCTGGCGAGCCAGCTAT	TCGGGCACATCGCTGCAACCACTTGACCC	6660
Db	6942	TCATATGCCAGCTCTCGGCTACCCAGCTGT	CCGCTCCATCTCTCAAGCACTTGACCC	7001
OY	6661	ACCACGCGCAACCCCTATATATGTGACAT	GTGTGAGTGAATGCTAACCT-----GTTCC	6708
Db	7002	GCCAAACCATATCTCCCCCTGACGCGCGAG	CTATATAGAGCTTACCTCTCTGTGAGGACGAG	7061
OY	6709	ATGGGGGGGCGATGTGACTCGGATPAGAGT	CTGGGTCCAAATGAGTGCCTTCTGAGACTCTC	6768
Db	7062	ATGGGGCGGCAACATCACACAGGGTTGAGT	CAGAGAACAAAGTGTGATTTCTGAGACTCTTC	7121
OY	6769	GACCCAATGTGTGAGAAAGAAAGACGCA	CTTGAACCTTTCGATACATCAGAAATATATGCTC	6828
Db	7122	GATCCGCTTGATGGAGAGGAGATGAGCG	GGAGGTCTCCGTAACCTGCAGAAATTCGCGG	7181
OY	6829	CCCAAGAAAGGTTTCCCAACAGCTTAT	CCGGCTGGGACACGCGCTGATTTCAACCCACCG	6888
Db	7182	AAGCTCTGGAGATTTCGCGGGCGCTCG	CGCTGTGGGCCCGGACCGGACTCAACCCCGCG	7241
OY	6889	CTTGTGGAAATCGTGAAGAAAGGCGCAG	ATTATCAACCGGCGCATGTGTCGGGGCTGTCTC	6948
Db	7242	CTATGTAGAGACGTGAAAGAAAGCTGAT	CTACGAACCACTGTGTCTCATATGCTTCCCGCTA	7301
OY	6949	CTCTCTCTTGAAGAAACCCCGACGCT	CTCCCGCCCAAGGAGCGCGGACAGTGGGCTTAAT	7008
Db	7302	CCACTCTCAAGGTCCCTCTGTGCTCT	CGCGCTCGGAAAAAGGTATCGGTGCTCTAC	7361
OY	7009	GAGGACTCTCAATAGAGATGCTCTTCA	CAACAGCTGTGCCATTAACTCTTTGGCAGCC	7068
Db	7362	GAATCAACCTTACTACTGTCTGTGG	CCGAGCTTGCACCAAAAGTTTGTGGCAGCTCTCA	7421
OY	7069	CCAAGCGGCGATTCAGGCTTTC	CCAAGGGGGGGGGCGTGGCCGATTCCGCGAGTCA	7128
Db	7422	ACTTCGCGCAATACGGGCGACAAAT	AGACAACTCTCTGAACCGCGC-----	7469
OY	7129	CTCTCTGATAGTTTGGCCCTTTTGGAGAC	GAGGTCCATCTTTCATATGCCCCCTCGAG	7188
Db	7470	CTTTCTGTGGTGGCCCCCGACCTCGAC	GTGAGTCTTATTTCTTCATATGCCCCCTCGAG	7529
OY	7189	GGGAGGTTGGAGATTCAGACTGAG	CGCTGAGCAGAGTATGAGCCCCCAACCCCCCCAG	7248
Db	7530	GGGGAGCTGGGAGATCCGGA-----		7549
OY	7249	GGGGGGGTGGCACTCCCGGCTCGAC	CTCGGGGTCTGTGTATCTTGTCTCGAGGAGAC	7308
Db	7550	-----TCTCAGCGACGGGTATAGT	TGACGCTATGATGAGGGGCGACACG	7595
OY	7309	GACTCCGTGTGTGTCTCTCATATGTAT	ACTCTCTGACCGGGGCTCTAATACTCTTGT	7368
Db	7596	GAAGATGTCTGTGTGTCTCAATGTCT	TATTTCTTGTGACAGGGGCACTGTGCATCCCGTGC	7655
OY	7369	AGTCCGAGAGGAGATTTACCGATTTAA	CCCCCTTGAGCACTCCCTGTTGCGATATAC	7428

Db 7556 GCTCGGGAAGAAACAAAACCTGCCATCAACGACGACGAACTGTTGCTACGCGCATCAC 7715
 Qy 7429 AACAGAGGTACTGTACACACAACAAGAGCGGCTCACTAAGGCTAAAGGTAACCTTT 7488
 Db 7716 AATCTGGTATTCACACACTTCACGAGTGTCTGCCAAGGCAAGAAAGTACATTT 7775
 Qy 7489 GATAGATGCAAGTGTCTGACTCTCTACTACGACTCACTTAAAGACATTAACTAGCTAG 7548
 Db 7776 GACAGACTGCAAGTTCGTGACAGGCACTTACAGAGACGTGCTCAAGAGAGTCAAGACGCG 7835
 Qy 7549 GCGTCAGAGGTCAACGCAAGGCTCTCTGCTAGAGGAAAGCTTCAGCTGACGCGCCCA 7608
 Db 7836 GCGTCAAAAGTGAAGGCTAACCTTGCTAGCAGGAAAGCTTCAGCTGACGCGCCCA 7895
 Qy 7609 CATTCGCAAGATCTAATATATGAGGCTTAAAGAGTCCGAGCTTTCGCGGAGG 7668
 Db 7896 CATTCAGCAAAATCCAACTTTGGCTATAGGCGCAAAAGCTCCGTTGCAATGCCAAG 7955
 Qy 7669 GCGCTTAAACAATCAAGTCCGTGTGAAGAGACCTCTGAGAGACTCAGAAACCAAT 7728
 Db 7956 GCGGTAGCCCAATCACTCCGTGTGAAGAGACCTCTGAGAGACAGTGTAAACCAATA 8015
 Qy 7729 CCCACAAACATTAATAGCCAAATAGAGTGTTCGCTGAGACCCCAACAGAGGCGCAAG 7788
 Db 8016 GACACTACATCATGAGCCAAAGAGGTTTTCGCTTCACTGAGAAAGGCGGCTCGT 8075
 Qy 7789 AAAGAGCTGCGCTTATGTTTACCTGAGCGGCTGAGAGCTGCGAGAGAGATGAGC 7848
 Db 8076 AAGCAGCTGCTCATGCTGTGTTCCCGACCTGGGCGTGGCGGTGTGCGAAGATGAGC 8135
 Qy 7849 CTTAATGACATTAACAACAACTTCTGAGCGGCTGATGAGGAGCTTCTTATGATTCAG 7908
 Db 8136 CTGTATGACGATGTATGACAGAGCTCCCTGGCGGTGATGAGAGCTCTTACGATTCAA 8195
 Qy 7909 TATTCGCCCGCTCAGCGGCTAGAGTGTCTTTGAAAGATAGGCGGAAAGAAAGACCTT 7968
 Db 8196 TACTACACAGGACAGCGGCTTGAATCTCTGTGCAAGGCTGAAGTCCAAAGAACCCG 8255
 Qy 7969 ATGGGTTTTGATGATACCGGATGCTTGAACCTCAACGCTCACTGAGAGAGATCAGG 8028
 Db 8256 ATGGGTTCTGTATGATACCGGCTTGTGAGGAGCTTTCACAGTCACTGAGAGGACATCCGT 8315
 Qy 8029 ACTGAGAGTCCATATATTCGAGGCTGCTCTTCCGAGGAGGCGCCACATGCCATAC 8088
 Db 8316 ACGAGGAGGCAATTTACCAATGTTGTGACCTGAGCCCCCAAGCCCGGTGCGCATCAG 8375
 Qy 8089 TCGCTAATCTGAGAGCTTATGCTGAGAGGCTATGTTCAACAGCAAGGCGCAAACTGC 8148
 Db 8376 TCCCTCACTGAGAGGCTTATGTTGAGGAGCTCTTACCAATTCAGAGGAGGAAACTGC 8435
 Qy 8149 GGGTACAGGCTTGGCGGCGCAGCGGAGTGTCTACACCTAGCATGAGGAAACACATCAG 8208
 Db 8436 GGGTACCGCAGGTGCGCGGCGCAGCGGCTATGACAACTAGCTGTGTGAACCTCTCAT 8495
 Qy 8209 TGTCTAGTGAAGCTTACGAGGCTTGTAAAGCTGACGAGATATTCGCGCCCAATGCTG 8268
 Db 8496 TGTCTACATCAAGGCGCGGAGCGCTGTGAGCGGAGGCTTCAGAGATGACCATGCTC 8555
 Qy 8259 GTATGCGGCGATGACTGTGTTGTCTACAGAAACCGAGGAGCCGAGAGAGAGAGAGG 8328
 Db 8556 GTGTGTGCGAGACTGTGTGTATCTGTGAAGATGCGGAGGCTCAAGAGAGAGCGGCG 8615
 Qy 8329 AACCTGAGAGCTTTCACGAGGCTATGACAGAGTATTTGCGCCCTCTGTGAGACCCGCC 8388
 Db 8616 AGCTGAGAGCTTTCACGAGGCTATGACAGAGTATTCGCGCCCTCTGTGAGAGCCGCC 8675
 Qy 8389 AGACCGAGATATGATCTGAGAGCTGATTAACATCTTCTCTCAATGTGTCTGTGCGCTG 8448
 Db 8676 CAACAGATATGAGCTTGTGAGCTTATTAACATCATCTCTCCAAAGTGTCAAGTCCGAC 8735
 Qy 8449 GGGCCACAAGCGCGCGAGATATCTGAGCAAGACCTTACCATCTGCATGCGCGG 8508

Db 8736 GACGCGCTGAAAGAGGCTTACTACCTTACCCGTCACCAACCCCGCTCGGAGA 8795
 Qy 8509 GCTGCTGGGAAACAGTATGACACTCCCTGTCATTTATGCTGGGAAACATCATCAG 8568
 Db 8796 GCGCGTGGAGAGACAGACAGACACTCCAGTCAATTTCTGTGAGGACATATATCAG 8855
 Qy 8569 TACGCGCGACATATAGGCTGCGATGCTCTGATGACACACTTCTTCCATTTCTATG 8628
 Db 8856 TTTGCCCCCACTGTGGCGAGGATGATATGATAGCCATTTCTTTAGCGTCTCATTA 8915
 Qy 8629 GCTCAGACAGCTGAGACAGAACTCAACTTTGATGATGATAGGAGCGGATGATCCGCG 8688
 Db 8916 GCGAGGATCAGCTTGAACAGGCTTAACTGTGATCTTACGAGACCTGCTACTCATTA 8975
 Qy 8689 AGTCCCTGAGCTCCACGCTATTAATGAAAGTTACATAGGCTTGAACGCTTTTCTCTG 8748
 Db 8976 GAACCACTGATTTACTTCATTCATTAATAAGCTCATAGGCTCAGGCAATTTTCACTC 9035
 Qy 8749 CACACATACACTCCCAAGACTGACAGCGGTGCTTCAAGCTTCAAGAAACTTGGGCG 8808
 Db 9036 CACAGTACTCTCAGGTGAATCAATAGGCTGCGCATGCTCAGAAAACTTGGGCTC 9095
 Qy 8809 CCAACCCCTGAGAGCGTGAAGAGCGGCGACGTCAGTACAGGCGCTCCATCTCCGT 8868
 Db 9096 CCGCCCTTGGAGCTTGAAGACACCGGAGCCGAGCGTCCGCTTGTCTGACAGA 9155
 Qy 8869 GGGGAGAGAGCGCGCTTGGCTGATATCTTCAATTTGAGCGGTGAAGACCAAGCTC 8928
 Db 9156 GAGAGGAGGCTGCTATATGTGCAAGTACCTTCACTGGGCAATTAAGAAACAAAGCTC 9215
 Qy 8929 AAATCACTCATTTGCCGAGCGCGCTCTGTGATTTATCCAGCTGTGTTACCGTCCG 8988
 Db 9216 AAATCACTCAATACCGCGCGCTGCGCTGAGCTGTGCTGTTGTTCAAGCGCTGCG 9275
 Qy 8989 GCGCGGAGGAGGACATTTATCAAGCGTGTGCGCGGAGCCCGCTTATGCTCTT 9048
 Db 9276 TACAGCGGAGAGACATTTATCAAGCGTGTGCTATGATCCCGCGCTGTCTGAGTT 9335
 Qy 9049 GGCCTACTCTACTTTTGTAGGAGTACGCTTTTCTACTCCCGCTCG 9098
 Db 9336 TGCCTACTCTGCTGCTGCGTGAAGGATGATGATCTCTCCCAACG 9385

RESULT 10
 AAC86648
 ID AAC86648 standard; DNA; 9611 BP.
 XX
 AC AAC86648;
 XX
 DT 02-APR-2001 (first entry)
 XX
 DE Nucleotide sequence of chimeric Hepatitis C virus clone pH77(p7)-J68.
 XX
 KW HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine; ss.
 XX
 OS Synthetic.
 OS Hepatitis C virus.
 XX
 FT Key location/Qualifiers
 FT CDS 342..9389
 FT /*tag= a
 XX
 PN WO200075338-A2.
 XX
 PD 14-DEC-2000.
 XX
 PF 02-JUN-2000; 2000MO-US015446.
 XX
 PR 04-JUN-1999; 99US-0137693P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 FI Yanagi M, Bukh J, Emerson SU, Purcell RH;

XX WPI: 2001-061728/07.
DR P-PSDB; AAB30733.
XX Nucleic acid molecule encoding human hepatitis C virus of genotype 2a for
PT developing vaccines, for diagnosis of hepatitis C virus and in screening
PT assays for identification of antiviral agents.
XX
PS Disclosure; Page 146-149; 167pp; English.
XX
XX AAC86645-48 represent chimeric cDNA clones of infectious Hepatitis C
CC virus (HCV). In each clone the C, E1 and E2 genes are derived from HCV
CC strain Hc-J5CH genotype 2a. The p7 protein was derived from HCV strain HC
CC -J6CH or HCV strain pcv-H77C, and the NS genes were derived from HCV
CC strain pcv-H77C. Such HCV sequences are capable of expressing the virus
CC when transfected into cells. The HCV protein is useful for assaying
CC candidate antiviral agents for activity against HCV. Antibodies specific
CC for HCV polypeptide are useful in prevention and treatment of diseases
CC caused by HCV in animals, in particular humans. The HCV polypeptides
CC serve as immunogens in the development of vaccines for preventing HCV in
CC mammals or as antigens in diagnostic assays for detecting the presence of
CC HCV in biological samples. The HCV polynucleotide is also useful for
CC identifying cell lines capable of supporting the replication of HCV in
CC vitro and to produce attenuated viral strains via passage in vitro or in
CC vivo
XX
XX Sequence 9611 BP; 1925 A; 2897 C; 2695 G; 2094 T; 0 U; 0 Other:
Query Match 58.0%; Score 5274.4; DB 5; Length 9611;
Best Local Similarity 74.1%; Pred. No. 0;
Matches 6751; Conservative 0; Mismatches 2281; Indels 78; Gaps 3;
QY 1 ATAGACCAAAATCTTAACCTCAAGAAAACCAAAAGAAACCAACCGTCCGCCCA 60
DB 342 ATAGACCAAAATCTTAACCTCAAGAAAACCAAAAGAAACCAACCGTCCGCCCA 401
QY 61 GACGTTAAGTTTCGGGCGCGGCGAGATGTTGGCGAGATATCTTGTGGCGGAGG 120
DB 402 GACGTTAAGTTTCGGGCGCGGCGAGATGTTGGCGAGATATCTTGTGGCGGAGG 461
QY 121 GGGCCAGAGTTGGGTGTGCGCGACAGAGAAAGACTTGGAGCGGTCCAGCCAGTGA 180
DB 462 GGGCCAGAGTTGGGTGTGCGCGACAGAGAAAGACTTGGAGCGGTCCAGCCAGTGA 521
QY 462 GGGCCAGAGTTGGGTGTGCGCGACAGAGAAAGACTTGGAGCGGTCCAGCCAGTGA 521
DB 462 GGGCCAGAGTTGGGTGTGCGCGACAGAGAAAGACTTGGAGCGGTCCAGCCAGTGA 521
QY 181 AGCGCCAGCGCCATTCCTTAAGATCGCGCTCCACTGGCAATCTGGGGAAAACAGGA 240
DB 522 AGCGCCAGCGCCATTCCTTAAGATCGCGCTCCACTGGCAATCTGGGGAAAACAGGA 581
QY 241 TACCCCTGGCCCTATACGGAATGAGGAGCTGGGCTGGGAGAGATGGCTCTGTCCCC 300
DB 582 TACCCCTGGCCCTATACGGAATGAGGAGCTGGGCTGGGAGAGATGGCTCTGTCCCC 641
QY 301 CGAGGTTCCGCTCTTGGGAGCCCAATGACCCCGGAGATAGTGCAGCACTGGGT 360
DB 642 CGAGGTTCCGCTCTTGGGAGCCCAATGACCCCGGAGATAGTGCAGCACTGGGT 701
QY 361 AAGGTATGATATCCCTTAAGATCGCGCTTGGCACTTCAATGGGATATCCCTGTGTG 420
DB 702 AAGGTATGATATCCCTTAAGATCGCGCTTGGCACTTCAATGGGATATCCCTGTGTG 761
QY 421 GGGGCCCCGCTGGGCGGCGTGCAGAGCTCTGGCGATGGCGTGAAGAGCTGGAGAG 480
DB 762 GGGGCCCCGCTGGGCGGCGTGCAGAGCTCTGGCGATGGCGTGAAGAGCTGGAGAG 821
QY 481 GGGGTTAATTTTCAACAGGAACTTAACCGGTTGCTCTTTCTATCTTCTTGGTGGC 540
DB 822 GGGGTTAATTTTCAACAGGAACTTAACCGGTTGCTCTTTCTATCTTCTTGGTGGC 881
QY 541 CTGCTGTCTGATCAACACCCGGGTCTCGCTGGCGAATGAAAGAACATCAGTACGGC 600
DB 882 CTGCTGTCTGATCAACACCCGGGTCTCGCTGGCGAATGAAAGAACATCAGTACGGC 941
QY 601 TACATGATGATCAACGACTGCACCAATGACAGATTAATCTGGAGCTCAGGCTGTG 660

DB 942 TACATGATGATCAACGACTGCACCAATGACAGATTAATCTGGAGCTCAGGCTGTG 1001
QY 661 CTCACGTCCTCCGGGTGCTCCGTGCGAAGATGGAGATCTCAAGTGTGATA 720
DB 1002 CTCACGTCCTCCGGGTGCTCCGTGCGAAGATGGAGATCTCAAGTGTGATA 1061
QY 721 CCGGTCAACGAAATGGGCGGTGACGCGGCGCGGCTTACGAGGGCTTGGGACG 780
DB 1062 CCGGTCAACGAAATGGGCGGTGACGCGGCGCGGCTTACGAGGGCTTGGGACG 1121
QY 781 CACATGACATGATGATGATGTCGCGCACCGCTGTCTGCTTCACTGAGGAGATCTC 840
DB 1122 CACATGACATGATGATGATGTCGCGCACCGCTGTCTGCTTCACTGAGGAGATCTC 1181
QY 841 TGGGTGGGTGATGCTCGAGCCCAATTTTCAATGTTCTGCGCGACACACTGTTT 900
DB 1182 TGGGTGGGTGATGCTCGAGCCCAATTTTCAATGTTCTGCGCGACACACTGTTT 1241
QY 901 GTCCAGACTGCAATGCTCATCTACCTGGTACATCACTGGACACCGCATGGCATGG 960
DB 1242 GTCCAGACTGCAATGCTCATCTACCTGGTACATCACTGGACACCGCATGGCATGG 1301
QY 961 GACATGATGATGAATGATGTCGCCACCGCTACCATGATCTTGGCGATGCGATGTC 1020
DB 1302 GACATGATGATGAATGATGTCGCCACCGCTACCATGATCTTGGCGATGCGATGTC 1361
QY 1021 CCGAGGTCAATTAAGACATCAATTAAGGCGGCTCATTTGGGCGTCAATGTTGGCCT 1080
DB 1362 CCGAGGTCAATTAAGACATCAATTAAGGCGGCTCATTTGGGCGTCAATGTTGGCCT 1421
QY 1081 TACTTCTTAATGAGGAGGCGTGGGAAAGTGGTATCTTCTGTGGCGCGCGG 1140
DB 1422 TACTTCTTAATGAGGAGGCGTGGGAAAGTGGTATCTTCTGTGGCGCGCGG 1481
QY 1141 GTGAGCGCGGCAACCAATCTGTGGGAGGTTCTGCGCGGAGACACCGGCGCTCAC 1200
DB 1482 GTGAGCGCGGCAACCAATCTGTGGGAGGTTCTGCGCGGAGACACCGGCGCTCAC 1541
QY 1201 AGCTTAATTTGACATGGGCGCCAGGACAGAAATCACTGTTAAACCAATGGCAGTGG 1260
DB 1542 AGCTTAATTTGACATGGGCGCCAGGACAGAAATCACTGTTAAACCAATGGCAGTGG 1601
QY 1261 CACATCAACCGCACCGGCTGAACTGCAATGATCTCTGACACCGGCTTATTCGCTCT 1320
DB 1602 CACATCAACCGCACCGGCTGAACTGCAATGATCTCTGACACCGGCTTATTCGCTCT 1661
QY 1321 CTGTTCTACACCAAGCTTCACTGTCAGATGTCGCCAAGCATGTCGCTGCGC 1380
DB 1662 CTGTTCTACACCAAGCTTCACTGTCAGATGTCGCCAAGCATGTCGCTGCGC 1721
QY 1381 AGTATGAGGCTTCCGGGTGGAGTGGGCGCTTGCATATGAGGATTAATGTCACCAAT 1440
DB 1722 AGTATGAGGCTTCCGGGTGGAGTGGGCGCTTGCATATGAGGATTAATGTCACCAAT 1781
QY 1441 CCAAGAGATTAAGACCTTAATGCTGGCACTACCAACCAAGAGTGGCGGTGCTCC 1500
DB 1782 CCAAGAGATTAAGACCTTAATGCTGGCACTACCAACCAAGAGTGGCGGTGCTCC 1841
QY 1501 GCGAAGACTGTGTGTGAGCCAGTGTACTGTTTCAACCCCGGCGAGTGTGAGGACG 1560
DB 1842 GCGAAGACTGTGTGTGAGCCAGTGTACTGTTTCAACCCCGGCGAGTGTGAGGACG 1901
QY 1561 ACCGAGAGGTTGAGCGCCCACTTAACAGTGGGAGAGATGAGACAGATGTTCTCTA 1620
DB 1902 ACCGAGAGGTTGAGCGCCCACTTAACAGTGGGAGAGATGAGACAGATGTTCTCTA 1961
QY 1621 TTGAACAGACTGAGCAACCGCTGGGGTCAATGTTGGCTGACAGTGGAGAACTCTCT 1680
DB 1962 TTGAACAGACTGAGCAACCGCTGGGGTCAATGTTGGCTGACAGTGGAGAACTCTCT 2021
QY 1681 GGTACACCAAGACTTGGCGCGCACCACTTCCGATTAAGAGCTGAATTCACGCGAGC 1740

Dh 2022 GCGTACACCAAGACTTGCGCGCACACCTCGCTAAGACTGACTTCAACGCGAC 2081
Qy 1741 ACGGACCTGTTGTGCCCCACGAGTGTGTTAGAAAGATCCTGATACACTTAACTCAAA 1800
Dh 2082 ACGGACCTGTTGTGCCCCACGAGTGTGTTAGAAAGATCCTGATACACTTAACTCAAA 2141
Qy 1801 TCGGCTCTGGGCGCTGCTCAAGCCAAAGTGCTGATCGAATAACCCCTCAAGGCTCTGG 1860
Dh 2142 TCGGCTCTGGGCGCTGCTCAAGCCAAAGTGCTGATCGAATAACCCCTCAAGGCTCTGG 2201
Qy 1861 CATTACCCCTGCAAGTTAATATATACATCTTCAAAATTAAGAGATGATGAGGAGGAGT 1320
Dh 2202 CATTACCCCTGCAAGTTAATATATACATCTTCAAAATTAAGAGATGATGAGGAGGAGT 2261
Qy 1921 GAGCACAAGGCTCAGCGCTGATGCAATTTCACTGAGGAGATCGTTGCAACTTGGAGAG 1980
Dh 2262 GAGCACAAGGCTCAGCGCTGATGCAATTTCACTGAGGAGATCGTTGCAACTTGGAGAG 2321
Qy 1981 AGAACAAGAGTCAACTGTCTCTTGTGTCATCTCCACACGAGATGAGGCAATTTAAGT 2040
Dh 2322 AGAACAAGAGTCAACTGTCTCTTGTGTCATCTCCACACGAGATGAGGCAATTTAAGT 2381
Qy 2041 TGCTCTTACTGAGACTGCGCGCTGTGTCATGCTGTCTTCCACTCCACCAAAATATC 2100
Dh 2382 TGCTCTTACTGAGACTGCGCGCTGTGTCATGCTGTCTTCCACTCCACCAAAATATC 2441
Qy 2101 GTGACAGTACAAATTCATGATGATGACTATACACTGCTCCACAAATATCATCTGCGAG 2160
Dh 2442 GTGACAGTACAAATTCATGATGATGACTATACACTGCTCCACAAATATCATCTGCGAG 2501
Qy 2161 GAGTGGGTAAATCTTATTTCTCTGCTCTTACGCGAAGCGCAGGTTTGCGCTGCTTATGG 2220
Dh 2502 GAGTGGGTAAATCTTATTTCTCTGCTCTTACGCGAAGCGCAGGTTTGCGCTGCTTATGG 2561
Qy 2221 ATGCTCATCTTGTGGGCGAGGCGAAGCAGCAGTACAGAAAGTGTGATCTTGCAGACT 2280
Dh 2562 ATGCTCATCTTGTGGGCGAGGCGAAGCAGCAGTACAGAAAGTGTGATCTTGCAGACT 2621
Qy 2281 GCGAGCGAGCTAGCTGCAATGCTCTCTATATTTGTGATCTTTTCTGAGCTGTGG 2340
Dh 2622 GCATCCCTGCGCGGAGCAGCGATCTTGTCCTTCCCTGCTGCTTCTGCTGCTTGG 2681
Qy 2341 TACATCAAGGCTCGGATAGTCCCTTACACTATTTCCCTGCACTGCGCTGTGCTCTT 2400
Dh 2682 TATCTGAAGGATGAGTGGGTCCCGAGCGGCTTACCGCTCTACGCGATGTGCGCTCTC 2741
Qy 2401 AGCTCATGCTCTAGCAATGCGCCCAAGGCTTATGCTTATGCGGATCTGTGATGCG 2460
Dh 2742 CTCTGCTCTGCTGCGCTGTGCTGCTGCGGCAATGCACTGACACGAGATGCGCG 2801
Qy 2461 CAGATAGGAGCGGCTGTGCTGATATGATCACTCTTTACTCTCACCCCGGATATAG 2520
Dh 2802 TCGTGTGGCGGCGTGTCTTGTGGGTTAATGCGGTGATCTGTGCGCATATTAAG 2861
Qy 2521 ACCCTTCTCAGCCGCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
Dh 2862 CGCTATATCAGCTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2921
Qy 2581 GTCCAGAGTGGGACCACTATGAGGTGCGCGGTGCGGTGATGAGATGAGGCG 2640
Dh 2922 CTGCACTGTGGGTTCCCTCCCTCAAGCTCCGCGGAGGCGCGCATGCGCTATCTACTC 2981
Qy 2641 GTCCGCAATTTCTACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2700
Dh 2982 ATGT 3041
Qy 2701 GGGCTGT 2760
Dh 3042 GAGCCCTTTGT 3101
Qy 2761 GCTCTACTGAGATGTGCAACATGCAAGCATCTCGCGGAGGCGAGGTACGTCAAGATG 2820
Dh 3102 GGCCTTCTCGGATGTGCGCGCTAGCGCGAAGATAGCGGAGGTCAATAGTCAATG 3161

Qy 2821 GCGCTACTAGCCCTTGGCAGGTGAGCTGCGACTTACATCTATGACCACTCAACCCCTATG 2880
Dh 3162 GCCATCATCAAGTATAGGCGCGCTTACTGCGACTATGTGTATTAACATCTCAACCCCTCTT 3221
Qy 2881 TCGGATTTGGGCTGTAGTGGCTGTGCGGACCTGCGGCTGTGCGGTTAGGCTTATCTTTC 2940
Dh 3222 CGAGACTGGGCGCACACGCGCTGTGCGAGATCTGGCGGTGTGTGTGTGTGTGTGTGTGT 3281
Qy 2941 AGTCCGATGAGAGAAAGTCAATTTGCTGGGAGCGGAGACAGCTGTGTGGGAGCATT 3000
Dh 3282 TCCGAAATGAGAACCAAGCTCATAGTGGGAGGAGATACCGCGCGTGTGTGTGTGTGTGT 3341
Qy 3001 TTAACAGACTTCCCGTGTCCGCGCACTTGTGTGGAGGTCTCTTGGCCCAAGTGTAT 3060
Dh 3342 ATCAACGCTTGGCCGCTCTGTGCGGTGAGGCGCAGAGATCTGTGTTGGGCGAGCGAC 3401
Qy 3061 GCGTATACTCCAAAGGAGT 3120
Dh 3402 GGAATGTCTCCAAAGGAGT 3461
Qy 3121 CGTGGCTTTGGGACCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180
Dh 3462 AGAGGCTCTGTAGGATGTATATACACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3521
Qy 3181 GGGGAAATTCAGGCTGT 3240
Dh 3522 GGTGAGGTCCAGATCGTGTCAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3581
Qy 3241 GTTTTGTGATGT 3300
Dh 3582 GTATGTGTGATGT 3641
Qy 3301 GTCAACCAATGTATCCAGT 3360
Dh 3642 GTCATCCAGATGTATCCAGT 3701
Qy 3361 ACTAAATCTTTGAGCGGT 3420
Dh 3702 TCCCGCTCATGTGACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3761
Qy 3421 GCTGATGTATCCCGGCTGT 3480
Dh 3762 GCGGATGTATCCCGGCTGT 3821
Qy 3481 CCTCTTTCACCTTGAAGGAGTCTCTCAGGAGGCGCGGTCTATAGCCCCAGGAGCCAGCT 3540
Dh 3822 CCATTTCTACTTGAAGGCTCTCTGCGGAGTCTCTGCGGAGTCTCTGCGGAGGAGCCAG 3881
Qy 3541 GTCGAGTCTTCCGGGACGT 3600
Dh 3882 GTGGGCTATTACAGGCGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3941
Qy 3601 CCGTGTGAACACTGCAATGT 3660
Dh 3942 CCGTGTGAAGACTGT 4001
Qy 3661 CCGT 3720
Dh 4002 CAGCAGT 4061
Qy 3721 AGCACAAGT 3780
Dh 4062 AGCACAAGT 4121
Qy 3781 TCAGT 3840
Dh 4122 TGT 4181
Qy 3841 AACATTAGGACTGT 3900
Dh 4182 AATATCAGGACCGGAGT 4241

OY	3901	GGCAAAATTCCTGGCGGANTGGGGGGGGTGGCGGGGGCGCGCTACGACATCATCATATATGGAT	3900
Db	4242	GGCAAGTTCCTTGGCCGAGCGGGGGCTCGAAGAGGTGTCTATATGACATATATATTTGGTAC	4301
OY	3961	GAATGGCATGCGCGTGAAGCTCTACCAACCATCTTGGGATCGGAAGATGCTTGTATCAAGCA	4020
Db	4302	GAGTGTCCACTCCAGAGATGCAACATCCATCTTGGGCAATGGCACTGTCTTGAACMAAGCA	4361
OY	4021	GAGACAGCTGGGGTCAAGCTAACTGTGTGCTCAAGCTACGCCCCCTGGGTCAGTACA	4080
Db	4362	GAGACTGGGGGGGAGACTGGTGTGTGTCTGCACTGCTACCTCCCGGGCTCGTCACT	4421
OY	4081	ACCCCCCAACCATPAGAGAGTGGCCCTTGGCAGAGAGGGCGAGATCCCCCTTCTAT	4140
Db	4422	GTGTCCACTCTTAATCAGAGAGGTGTCTGTTCACACCGGAGAGATCCCCCTTTAC	4481
OY	4141	GGGAGGGGCGATTCCTCCCTGTCTTAACATCAAGGAGAGAGACATCTGTATCTTGCATTTCA	4200
Db	4482	GGCAAGGCTATCCCTCCCTCGAGGTATCAAGGGGGAGACATCTCATCTTCTTGCACCTCA	4541
OY	4201	AAGAAAAAGTGTGACGAGCTGCGGGCGGCCCTTGGGGTATGGGCTTGAATCTAGTGGCA	4260
Db	4542	AAGAAAGAGTGCAGAGAGCTGCGCGGCAAGCTGGTGTGATTTGGGATCAATGCCGTGGCC	4601
OY	4261	TACTACAGAGGTTGAGAGCTCCGTATATCAAACTCAGGGAGAGTATGTGTGTGCC	4320
Db	4602	TACTACCGCGGTCTTTCAGTGTCTGTATCCCGACACAGCGCGAGTGTGTGTGTGTGCG	4661
OY	4321	ACCGAGCGCCCTCATAGACAGGGTATATCTGGGGGACTTTGACTCGAGATTCAGTGAAGTA	4380
Db	4662	ACCGATGCTCTCATGACTGGGCTTTACCGGGCACTTTCGCTTGTATATGATGTGCACACG	4721
OY	4381	GCGGTCACTCAAGTGTATGACTTTCAGTTTGAACCCCACTTCAACATTAACACACACAGTT	4440
Db	4722	TGTGTCACTGAGACAGTGTGATTTACGCTTATACCTTACCTTATGAGACACACAGC	4781
OY	4441	GTCCTCTAAGACGCTGTCTACAGTACCAAGCGCGGGGTGCACAGGGTAGGGAGAACTG	4500
Db	4782	CTCCCCCGAGTGTGTCTCCAGGACTCAACGCGGGGCAAGACTGGCAGGGGGAGAACCA	4841
OY	4501	GGCAATTAATAGTATNGTTTTCACCTGTAGAGGAGACCTCAGGAATGTTTGAACAGTATGTG	4560
Db	4842	GGCATCTATATATTTTGTGGCACCGGGGAGGGCCCTTCGGCAATGTTCAGTCTGTCCGTC	4901
OY	4561	CTCTGTGAGTGTCTACGACGACGAGGGGCGCATGTATGAGCTCACACCATCGGAGACACC	4620
Db	4902	CTCTGTAGTGTCTATGAGCGCGGGCTGTCTTGTGTATGAGCTCACGCCCGCAGAACATCA	4961
OY	4621	GTCAGGCTCAGGGGCGTATTTTCMAACGCGCGGGTTCCTGTGTGCCAAGACCATCTTGAG	4680
Db	4962	GTTAGGCTACGAGGCTACATGAAACCCCGGGGCTTCCGTTGTCAGAGACATCTTGAA	5021
OY	4681	TTTTTGGAGGCAAGTTTTCACCGGCTCACACATATATGCGCATCTTCTTTCCMAACA	4740
Db	5022	TTTTTGGAGGGCGTCTTTTACGGGGCTCACTCATATATATGTCACATTTTATATCCAAACA	5081
OY	4741	AAGCAATCGGGGAAAAATTTTGCATATCTTAAACAGCTTACAGGCTTACAGGCTTACAGTGG	4800
Db	5082	AAGCAGATGGGGAGAACTTTCTTCACTGTGTAGCGTACCAAGACCAACGTTGTGGCTAGG	5141
OY	4801	GCCAAAGCCCCCCCCGCTCTGTGGACGTCAATGTGGAATGTTTGAATCTGCACCTCAAGCC	4860
Db	5142	GCTCAAGGCCCTCCCCCATCGTGGGACCAAGATGTGAAAGTGTTTGATTCGCTTAAACC	5201
OY	4861	ACACTCGTGGGCCCCCAACCTCTCTGTACCGCTTTGGGCTTGTTTACCAACAGAGTCAAC	4920
Db	5202	ACCTCTCATGGGCAACACCCCTGTATACAGACTGGGCGCTGTTCAGAAATGAAGTCAAC	5261
OY	4921	CTCACACATCCCGTGAAGAAATATATCGCCACTGTACATGCAAGCGCACTTGAAGTCAATG	4980
Db	5262	CTGACGACCCAAATCAACAAATATCATATGACATGTATGTGGCGCCAGACTGTGAAGTGTCT	5321
OY	4981	ACCAAGCATATGGTCTTGTGGCAGGGGAGCTTGTGGCGCGCTGTGCTGATTTGCTGGCG	5040

Dp	5322	ACGAGCACCTGGGTGCTCGTTGGGCGGCTCTGGCTGCTCGGCGCGATTCGCTTCA	5381
Qy	5041	ACCGGATGTGTTTGTCATCATCGGCGCTTGCACTTAACAGCAGCGCTGTTGCCCG	5100
Dp	5382	ACAGGCTGGCGGTCATATGATGGGAGAGATGCTGTGTCGGGAAAGCCGGCAATTATACCT	5441
Qy	5101	GACAGAAGAGTCTCTATGAGGCTTTGATGATGAGAGATGTGCTCTTAGGGGCGGCT	5160
Dp	5442	GACAGGAGAGTCTCTACAGAGGTTGATGATGAGAGAGTGCTCTCAGCACTTACCG	5501
Qy	5161	CTCATTAAGAGGGGGACGGGATPAGCGGATAGTGTAAGTCCAGATCCAAAGCTTATAG	5220
Dp	5502	TACATCGAGCAGGAGGATGATGCTTCGTCGAGCGAGTTCAAGCAGAAAGCCCTCGGCTCTTG	5561
Qy	5221	CAGCAAGCTTCCAAACAGGCTCAAGACATACAAACCACTGTGCAAGGCTTACATGCGCAG	5280
Dp	5562	CAGACCGCGTCCCGCCATGCAAGAGTATACACCCTGCTGTCCAGACCAACTGGCAAGAA	5621
Qy	5281	GTAGAACATTTCTGGGCGAAACACATGTGAACTTCAATAGCGGCATCCATACCTCGCA	5340
Dp	5622	CTCAGAGTCTTTTGGGCGAAGACATGTGAAATTTCACTAGTGGGATACAAATATTGGCG	5681
Qy	5341	GGAATATCAACATCTGCCAGGGAAACCTTCGAGTAGCTTCCATGATGGGTTCAGTGCCGC	5400
Dp	5682	GGCCTGTCAACGCTGCTGTGTAACCCCGCAATGCTTCATGTATGCTTTTACAGCTGCC	5741
Qy	5401	CTCACAGTCCGCTGTCAACAAAGCAGCACTATCCCTTCTCAACATTTTGGGGGCGTGCTA	5460
Dp	5742	GTACACAGCCCACTAACCACTGGCCAAACCTCTCTTCAACATATTTGGGGGGTGGGCTG	5801
Qy	5461	GCATCCCAAAATYGACCAACCCGCGGGGCGCATGCGCTTGCTGTGATGGCTTAATGAGGA	5520
Dp	5802	GCTGCCAGCTCGCGCGCCCGGGGCGGCTACTGACCTTTTGAGGATGCTGGCGTATGCTGC	5861
Qy	5521	GCTGCCGTAGGACGTAATAGCTTAAGTAAGTGTATGGAACATCTTCGCGAGGTTATGCT	5580
Dp	5862	GCGCCCATCGGACAGCTGTGAGCTGGGGAAAGTCTCTGTGACATTTCTTGACAGGGTATAGC	5921
Qy	5581	GCGGGCAATTTGCGGGGGCTCTGTCGCAATTCAGATCATGTCGGCGAAGACCTCCATG	5640
Dp	5922	GCGGGCGTGGCGGGAGCTCTTTAGCATTTCAAGTATGAGCGGTAGAGTCCCTCCACG	5981
Qy	5641	GAGGATGTCGTCAACTTGTGCTGCTGGAATTCGTCTTCGCGGATGCTTGTGATGGAGTCT	5700
Dp	5982	GAGGACCTGTGCAATCTGCTGCCGCGCATCTCTCGCTGGAAGCCTTTGATAGTGGGTG	6041
Qy	5701	ATCTGCGCGGCATTCTGCGCCGACACGTGGAGACCGGGGGAAGCGCCGTCCAATGAGATG	5760
Dp	6042	GTCGCGCACCAATATCTGCCGCGCAGTTGGCCGCGGCGACAGGGGAGCATGTAGATG	6101
Qy	5761	AATATGACTCATTTGCTTGTCTTCCAGAGAAATACGATGCGCCCAACCACTACGTAGAC	5820
Dp	6102	AACCGGCTAATAGCTTTCGCTCCCGGGGGAACATGTATTTCCCAAGCACTAAGTCCG	6161
Qy	5821	GAGTCGAGTACGATGCGAGCGTGTGACCCCACTACTTGTGCTCCCTTACATTAACAGCGTG	5880
Dp	6162	GAGAGCGATCAGCGCGCCCGCTGCACTGCAATCTCAGCAGCGCTCACTGTATACCCAGCTC	6221
Qy	5881	CTCAGAAAGACTCCACAATGTGATTACTGAGAGCTGCCCATGCCATGCGGCGGCTGCTGG	5940
Dp	6222	CTGAGGCGGACTGCTACGTGGAATGACTCGGAGGTGTACACTCCATCTCCGCTTCTCGG	6281
Qy	5941	CTCCGCGATGTGTGGGACTGGGTTTGGACATCTCTAACACTTTTAAAAATTTGGCTAGC	6000
Dp	6282	CTTAAGGGAATCTGGGACTGGAATATGAGAGTGTGAGCACTTTTAAGACTTGCTGATAA	6341
Qy	6001	TCCAAATTAATTTCCCAAGATGCCCGGCTCCCTTGTCTCTGTCAAAAAGGGGTACAG	6060
Dp	6342	GCCAAAGCTTAATGCCAACAATGCGCTGGGATTCCTCTTGTGTCTGTGCCAGCGGGGTATAGG	6401
Qy	6061	GCGCTGTGGCGCGCACTGGCAATCATGACCAACAGGTGTCTTTCGCGCGCAATATCTCT	6120

D 6402 GGGGCTGGCAGAGAGACGCAATTAATGCACCTGCTGCACTGTGAGCTGAGATCACT 6461
OY GGCATATGCTCGCTTGGGCTTCCATGAGATCAAGGGGCTTAAGACCTGATGAATATCTGG 6180
D 6462 GGACATGTCAAAAGGAGGACATGAGATGCTGGTCTTAGAGCTTGAGAAACATGTGG 6521
OY 6181 CAGGGAGCTTTCCTATCAATTTGTTTACAGGAGGCGCAGTGGCCGAAACCCGCGCA 6240
D 6522 AGTGGAGAGTTCCTCATTAACGCTACACAGGGGCTCTGACTCCCTTCCTGCGCG 6581
OY 6241 AACTTAAAGTTCGCACTGAGAGGGTGGCGGCTCAAGATACGGGAGGTGACCCAGAC 6300
D 6582 AACTATAAGTTCGCGCTGTGAGGGGTGTCTCAAGAGATACGTGAGATATAAGCGGGTG 6641
OY 6301 GGGTCATACCACTACATACAGACTCACACTGATTAATTGAAGTCCCTGCGCACTA 6360
D 6642 GGGGACTTCACATAGATATCGGGTATGACTACTGACAACTTTAAATGCGCTGCGACATC 6701
OY 6361 CCTCTCCGAGTTCTTTCTGGGTGAGCGGAGTGCAGATCAATAGTTTGGCCGACA 6420
D 6702 CCATGCGCCGAATTTTTCACAGAAATTTGACGGGGTGCGCTACACAGGTTTGGCCCT 6761
OY 6421 CCGAAGCCGTTTTTCCGGAGTGAAGTCTGCTTGGCTTGGGCTTAATTCATTTGTGTC 6480
D 6762 TGCAGAGCTTGTGCGGAGAGAGATCATTCAGAGTACAGCTCAAGATACCCGGTG 6821
OY 6481 GGGTCCAGCTTCTTGGCAGCCCTGAACCCGACACAGACGATTTGATGCTCATGA 6540
D 6822 GGGTGCAGATTAACCTTGAGAGCCGACCGAGCGTACGCGGTGGTGGCTCATCTACT 6881
OY 6541 GATTCATCTATATCAGCGGAGAGCTGACGCGGCGCTTTAGCGGGGTCAACCCCA 6600
D 6882 GATCCCTCCCATATPACAGAGAGGCGGCGGAGAGAGTTGGGAGAGGGTCAACCCCT 6941
OY 6601 TCCGAGGAGAGTCTCTCGGAGAGCCAGCTATCGGACATGCGTGGAGCCACCTGACCC 6660
D 6942 TCTATGCGAGTCTCTCGGCTAGCGAGCTGTCTGCTCATCTCTCAAGGCAACTTGGACC 7001
OY 6661 ACCCAGGCAAAAGCTATGATGTGAGCATGTGATGTCTAACT-----GTTG 6708
D 7002 GCCAACCATGACTCCCTGACGCGGAGCTCATAGAGCTTAACCTCTGTGAGAGCGAAG 7061
OY 6709 ATGGGGGCGATGTGACTCGATAGAGTCTGGGTCAAAAGTGTGTTCTGGAATCTCTC 6768
D 7062 ATGGCGGCAACATACACGAGGTTGATCAGAGAACAAAAGTGTGATCTGGAATCTCTC 7121
OY 6769 GACCCAAATGTGTGAGAAAGAGGACCTTGAAGCTTGTGATACCATGAGATACATGTC 6828
D 7122 GATCGGCTTGTGAGAGAGAGATGAGCGGAGGTCTCGTACCTGCAAGAAATTCGTGCG 7181
OY 6829 CCAAGAGAGAGTTCACACAGCTTTACCGGCTTGGGACGCGCTGATTAACAACCAACG 6888
D 7182 AAGTTCGAGAGATTCGCGCGGCTTGCCTGCTGTGGGCGCGGAGCTTACAAACCCCGC 7241
OY 6889 CTTGTGAGATCGTGAAGAAAGCGCAGATTAACAACCGGCACTGTTGCGGGCTGTGCTC 6948
D 7242 CTAGTAGAGAGCGTGAAGAAAGCTGATACGAACCACTGTGTGTCATGAGCTGACCGCTA 7301
OY 6949 CCTCTCTTAAGAAAAACCCGACGCTCTCCCAAGAGAGCGCGGACAGTGGGCTTAAGT 7008
D 7302 CCACCTCACAGGTCTCTCTGTGTGCTCGCTCGGAAAAAGCGTACGCTGCTCTCAAC 7361
OY 7009 GAGGACTGCATGAGATGCGCTTCAACAGCTGGCCTTAAGTCTTTTGGCAGGCCCCC 7068
D 7362 GAATCAACCTTATCTATGCTTGTGGCGAGGCTTGCCAAAGTTTGGAGCTCTCA 7421
OY 7069 CCAAGCGGCGATTCAGGCTTTTCAAGGGGCGGGCGCTGCGATTCGCGAGTCAAGCG 7128
D 7422 ACTTCGGGCAATTAACGGGCAACAATACGAACAATCTCTGAGCCCGC----- 7469
OY 7129 CCTCTGTAGAGTGGGCTTTTGGAGACAGGTTTCAATCTCTTCAATGCCCCCTCGAG 7188
D 7470 CTTCTGAGCTGCCCCCGAGCTCGACGTTGAGTCTTATCTTCATGCCCCCTCGAG 7529

OY 7189 GGGAGCTTGAAGATCCAGACTGAGGCTGAGAGAGTAGAGCCCAACCCCCCCAG 7248
D 7530 GGGGAGCTTGGAGATCCGA----- 7549
OY 7249 GGGGGGATGAGACTCCCGCTCGAGCTGGGCTCTGATCTTATCTTCCAGAGAGAC 7308
D 7550 -----TCTCAGCGACGAGGTATGCTACCGGTGATGAGTGGGCGCGACAG 7595
OY 7309 GACTCGGTGTGTGCTCATATGCTATCTCTGTGACCGGGGCTTAATTAATCTCTGT 7368
D 7596 GAGAGTGTGTGTGCTCATATGTTTATCTTGTGACAGGGGCACTGTCAACCCGTGC 7655
OY 7369 AGTCCGAAAGAGAGATTTACCGATTAACCCCTTGGAGCACTCCCTGTTGGATATCAC 7428
D 7656 GCTGGGAGAGAAACAAAACCTGCCATCAAGCACTGAGCACTGTTGCTACGCAATCAC 7715
OY 7429 AACAGGTGTACTGTACCAACAAAGAGGCTCACTAAGGCTTAAGGTAAGTAACTTTT 7488
D 7716 AATCTGTGTATTCACACTTCACGAGTGTGCTCAAGGCGAAGAAAGTCACTATT 7775
OY 7489 GATAGATGCAAGTGTGCTCTACTCTCACTCACTCACTCACTCACTCACTCACTCACT 7548
D 7776 GACAGACTGCAAGTGTGAGCAGCTTACAGAGAGCTGCTCAAGAGGTCAAGAGAGCG 7835
OY 7549 GCCTCAAGGTCAACCGCAAGGCTCTCACCATTGAGAGAGGCTTGCAGTTAAACCCACCC 7608
D 7836 GCGTCAAAAGTGAAGCTTAATCTGATCGTGAAGAGAGCTTGCAGCGCTGACGCCCA 7895
OY 7609 CATTCTGCAAGATCTAAATATGAGGTTTGGGGCTTAAGAGGTCCGACGTTTTCGGAGAG 7668
D 7896 CATTCAAGCAAAATCCAAAGTTTGGCTATGGGGCAAAAGAGCTCGTTGCCATGCAAAAG 7955
OY 7669 GCGGTTAACCAATCAAGTCCGTGTGAGAGAGCTCTGAGAGAGCTCAGAAACCAAT 7728
D 7956 GCGTGAACCAATCAATCTCGTGTGAGAGAGCTTCTGAGAGAGCTTACACCAAT 8015
OY 7729 CCCCAACCAATTAAGCCAAAAATGAGTGTCTGTGAGACCCCAACAAAGGGGCAAG 7788
D 8016 GACATACCATCATGCGCAAGAGAGGTTTCTGCGTTCAAGCTGAGAAAGGGGCTCGT 8075
OY 7789 AAGCAGCTGCGCTTATCGTTTACCTGACCTCGGCTGACAGGCTCTGCGAAGATGCGC 7848
D 8076 AAGCAGCTGCTCATCTGTTTCCCGAGCTGGGGGTGCGCTGTGCGAGAAAGTGGCG 8135
OY 7849 CTTTAGACATACAAAAACCTTCAGGCGGATAGGGGCTTCTTATGATTCAG 7908
D 8136 CTGTAGAGCTGTGAGAGCTCCCTGCGCTGAGCTGAGAGAGCTCTTACGATTTCAA 8195
OY 7909 TATTCCCGCTCAGCGGGTAGAGTTCCTTGAAGAGATGGCGGAAAGAGAGCCCT 7968
D 8196 TACTCAGCAGGAGAGGGGGTTGAATTCCTGTGCAAGCGTGAAGTCCAAGAGACCCG 8255
OY 7969 ATGGGTTTTGTATGATACCGGATGCTTTGACTCAACGCTCACTAGAGAGATCAG 8028
D 8256 ATGGGGTTCGTATATATCCCGCTGTGTTGATCTCCACAGTCACTAGAGCGACATCCG 8315
OY 8029 ACTGAGGATCATATATGCGGCTGCTCTTCCGCGGAGAGGCGCACACTGSCATAC 8088
D 8316 ACGAGGAGGCAATTAACATATGTTGACCTGAGACCCCAAGCCGCTGSCATACAG 8375
OY 8089 TCGCTAATGAGAGACTTTACGTGAGAGGCGCTTATGTTCAACAGCAAGGCGCAACCTGC 8148
D 8376 TCCCTCACTGAGAGGCTTTATGTTGGGGGCTCTTAACAAATTCAGAGGGGAGAAATGCG 8435
OY 8149 GGGTACAGCGCTTGGCGCGCAGCGGGGTCTCAACATAGCATGAGGAGAACATCATCA 8208
D 8436 GGGTACCGCAGGTGGCGGAGCGGAGCTGATCAACAATAGCTGTGTAAACCTTCATC 8495
OY 8209 TGCATGATGAAGCTTAAAGGCTTGAAGAGTGCAGGAGTAATTCGCGCCCAATATGCT 8268
D 8496 TGCTCATCAAGGCGCGGAGCTGTGAGCGGAGCTTCAAGAGCTGACCATGCTC 8555

OY	8269	GTATGCGGCGATGACTTGTTGGTGTGATCTCACAAGAACCGAGGGGACCGAGAGAGACGAGCGG	83228
Db	8556	GTGTGTGCGACGACTTATGTCGTATTCGTGAAGTGTGCGGAGGTCACAGAGACGCGGCG	86155
OY	8329	AACCTGAAAGCCTTCAACGGAAGCTATGACACAGATATTTGACCCCTCCTGTGACCCCCC	83888
Db	8616	AGCCTGAAAGCCTTCAACGGAAGCTATGACAGATATTTGACCCCTCCTGTGACCCCCC	86757
OY	8389	AGACCGAGTATGATCTGAGCTGATTAACATCTTGCTCTCAATATGTGTCTGTGGCGCTG	84488
Db	8676	CAACGAGAAATCGACTTGGAGCTTATTAATATCATGCTCTCCAAAGTGTGAGTGCAC	87353
OY	8449	GGCCCAACAAGGCGCGCGAGATPACTACCTGACAGAGACCCCTACACTCCAAATGACCCCG	85088
Db	8736	GACGCGCTGAAAAGAGGGGTCTACTACTTACCTGACCCGTGACCTTACAAACCCCTCGCGAGA	87955
OY	8509	GCTGCTGTGGAAAGATTAAGACACTCCCTCTGTCAATTCATGAGTGGGAAATCATCATCAG	85688
Db	8796	GCGCGTGTGGAGACAGCAAGACACTCTCACTCAATTCCTGTGATGGCAATATCATG	88555
OY	8569	TACGCGCCGACATATGAGCTGTGCAATGCTGTGATGACACACTTCTTCTCAATTCATG	86288
Db	8856	TTTGCCTCCCACTGTGTGGCGAGATGATATCTGATGACCACTTCTTAAAGGTCTCTATA	89155
OY	8629	GCTCAAGACAGCTGTGACCAAGAACTCAACTTTGATGATGACGAGCGGTGTACTCCGTG	86888
Db	8916	GCCAGGATTCAGCTTGAACAGGCTTAACTGTGTGATCTACGAGAGCTGTCTACTCTATA	89755
OY	8689	AGTCCCTTGGACCTCCCGAGCTATPATAATGAAAGTTACAATGGGCTTGAACGCTTTTCTCTG	87488
Db	8976	GAACTACTGATCTTACTCTCAATCTTAAAGACTCCATGTGCTCAAGCGCATTTTCACTC	90355
OY	8749	CACACATTAACCTCCCAAGAACTGACACAGGTGTGCTTCAAGCCTCAGAAAATTGAGCG	88088
Db	9036	CACAGTTACTCTCCAGGTGAATATCAATAGGGTGGCGGATGCTCTAGAAAATTGAGCGT	90955
OY	8809	CCACCCCTCAGAGCGTGAAGAGCGCGGACAGTGAAGTCAAGGCGTCCCTCATCTCCGT	88688
Db	9096	CCGCGCTTGGCAGGCTTGGAGACACCCGCGGCCGAGAGCGTCCGCTAAGGCTTCTGTCCAGA	91555
OY	8869	GGGGGAGAGCGGCGCTTGGCGTTCGATATCTTCTCAATTTGGGCGGTGAACAACAAGTCT	89288
Db	9156	GGAGGACAGGGGTGTATATGTGGCAAGTACTCTTCAACGTGGGACATGAAGAAACAAGTCT	92155
OY	8929	AAATCTCACTCATTTGCGGAAAGCGCGCTCTGTGATTTATCAGCTGTTCACAGTCCGC	89888
Db	9216	AAATCTCACTCAATAGCGCGCGCTGTGCGGCTGTGACCTTGTCCGATTCACGCGTGC	92755
OY	8989	GCCGCGGAGGAGCAATTTATCAGACGCTGTGCGGTGCCGAGCCCGCTTATTTCTCTT	90488
Db	9276	TACAGCGGGGAGACATTTATCAGACGCTGTCTCATGCTCCGCGCCCGCTGTCTCGGTT	93355
OY	9049	GGCTTACTCTTCTTTTGTATGGGGTGAAGCCTTTTCTTCACTCCCGCTCG	90988
Db	9336	TGCCTACTCTGTCTGCTGACAGGGGTAGGACATCTACTCTTCCCAACCG	93855
RESULT 11			
AA063499			
ID	AA063499	standard; cDNA; 9436 BP.	
XX	AA063499;		
AC			
XX	17-JAN-1995	(first entry)	
DT			
XX		Blood transmissible NANBHV genome.	
DE			
XX			
KW		Polymerase chain reaction; PCR; amplify; primer; non-A, non-B hepatitis;	
KW		NANBH; virus; blood transmissible; detection; hepatitis virus; RT-PCR;	
KW		C100 antibody; HCV RNA; NS5 region; ds.	
XX			
OS	Non-A.		
OS	non-B hepatitis virus.		

	Key	Location/Qualifiers
XX FH CDS	342..9374	
XX FT	/tag= a	
XX PN	JP06105690-A.	
XX PD	19-APR-1994.	
XX PF	10-MAR-1992; 92JP-00051885.	
XX PR	10-MAR-1992; 92JP-00051885.	
XX PA	(KAEN/) KAENNO K.	
XX DR	WPI; 1994-163130/20.	
XX DR	P-PSDB; AAR53417.	
XX PT	Blood-transmissible non-A non-B hepatitis virus DNA - used for detection	
XX PT	of hepatitis virus.	
XX PS	Claim 1; Page 8-20; 22pp; Japanese.	
CC CC	This sequence represents the genome of a blood transmissible non-A non-B	
CC CC	hepatitis (NANBH) virus. This sequence was isolated using the primers	
CC CC	given in AAQ63500-35. The amplified fragments are used in the detection	
CC CC	of hepatitis virus. This target DNA was isolated from serum of	
CC CC	chronically infected NANBH patients who were C100 antibody-positive and	
CC CC	HCV RNA (NS5 region) positive. Reverse transcription-PCR and PCR were	
CC CC	performed on cDNA and the total human NANBH DNA was constructed from 23	
CC CC	clones	
SQ	Sequence 9436 BP; 1876 A; 2840 C; 2655 G; 1974 T; 0 U; 91 Other;	
Query Match	45.7%; Score 4160.2; DB 2; Length 9436;	
Best Local Similarity	66.2%; Pred. No. 0; Mismatches 2924; Indels 93; Gaps 6;	
Matches 6029; Conservative	65;	
OY	1 ATGAGCAACAATCCTTAAACCTCAAGAAGAAAAACAAAAGAAACACCAACCCTGCCACAA 60	
Db	342 ATGAGCAACAATCCTTAAACCTCAAGAAGAAAAACAAAAGTAACACTACCGCCGCACAG 401	
OY	61 GACGTAAAGTTTCGGGCGGCGGCAGATCGTTGGCGAATACTTGTCGCGCAGG 120	
Db	402 GACGTCAAAGTTCGGGCGGCGGTGTCAGATCGTTGGTAAGTTAACGTGTTCCGCGCAGG 461	
OY	121 GGCCCCAGATTGGGTGTGCGCGCACGAAGAACTTCGAGACGGTCCACGCCATGTGA 180	
Db	462 GGCCCCAGTTGGGTGTGCGCGCACATGAGAACTTCGAAACGATGCGCAACTGTGGA 521	
OY	181 AGGCGCAGCCCATCCCTTAAAGATCGGCGCTCCACTGCAAACTCTGGGAGAAACAGGA 240	
Db	522 AGGCGAACAACCTTATCCCAAGGCTCGCCGACCCAGAGGTAGGGCTGAGCTCAGCCTGGG 581	
OY	241 TAACCCCTGACCCTTAAACGGAATGAGGAGCTCGGCTGGGCGAGATGGCTCTGTCGCC 300	
Db	582 TACCTTAGGCCCTCTATATGGAATGAGGAGCTGGGGTGGGAGGATGAGCTCTGTACCC 641	
OY	301 CGAGTTTCGGTCCCTCTTGGGGGCCCAATGACCCCCCGGCAATAGTGGCGCAACGTGGGT 360	
Db	642 CGCGGCTCCCGGCTTAGTTGGGGGCCCAACGACCCCAGGAGGTGCGCGAACTTGGGT 701	
OY	361 AAGGTCAATGATACCTTAAACGTGGGCTTTGGCGACCTCAATGGGGTACATCCCTGTGCTG 420	
Db	702 AAGGTCAATGATACCTTCAAGTGGGCTTTGGCGACCTCATGGGTACATTCGCTGCTC 761	
OY	421 GGCGCCCCGCTCGCGGCGTGCACGACTTTCGCGCATGGCGTGAAGTCTTGAAGAC 480	
Db	762 GGCGCCCCCTAAGGGGGGCGTGCACAGGAGCCCTGSCAATGGGTTCGGAAGAC 821	
OY	481 GGGGTAAATTTTGCACAAGGAACCTTACCGGTTGTCTTTTATCTTTCGCTGGCC 540	
Db	822 GGGGTAAATTTTGCACAAGGAACCTTACCGGTTGTCTTTCTATCTTCTCTTGGCC 881	

OY 541 CTGCTCTGATCAACCCCGGCTCCGCTGCCGAAGGAACATGATCCGCG 600
| | | | |
DB 882 CTGCTCTCTTTTGAACATCCGAGCTTCGCTTATCAGGTGCCGAAGCATCCGGGCTG 941
| | | | |
OY 601 TACATGTGATTAACGACTGCAACAAATGACAGCATTTACTGACGCTCCAGCTGCTGTC 660
| | | | |
DB 942 TACCATGTACGAATGACTGCTCCAACTCAAGCATTTGTGTATGAGGACGGGGCATGATC 1001
| | | | |
OY 661 CTCACGCTCCCGGGTGGTCCCGTGGAGAAAGTGGGGAATGATCTCAAGTGTGATA 720
| | | | |
DB 1002 ATGCTTACCCCGGGGCGTACCTGTGCTTGGAGAAACAAGCTCCCGTGTGGGTR 1061
| | | | |
OY 721 CCGGCTCTACCGAATGTGGCCGTGCGAGCGGCCCGCGCTCAGCGAGGCTTGCGAGC 780
| | | | |
DB 1062 GGGCTACTCCCAAGTGTGGCGGCCAGAAACAAGCATCCCACTAGACAAATAGAGCC 1121
| | | | |
OY 781 CACATGCAATGTGTGTATGTCCGCCACGCTGTCTGTCCCTCTACGTGGGGAGCTTC 840
| | | | |
DB 1122 CACGTCTGACTTGTGCTGTGGGGCGGCTCTTCTGTCCGCCATGTACGTGGGGAGCTTC 1181
| | | | |
OY 841 TGGCGTGGGGTGAATGCTGGAGGCCGAATGTTCATTGTCTGGCGGACAGCACTGCTTT 900
| | | | |
DB 1182 TGGGATCCGTTTCTCTGCTCTCCAGTTGTACCTTCTGCTGCGCGGTATGAGACA 1241
| | | | |
OY 901 GTCCAAGACTGCAATTGCTCCATCTACCTGTGTACCATCTGAGACACCGCATGSCATGG 960
| | | | |
DB 1242 GTACAGGACTGGAATTTGCTCAATCTATCCGACCAATGATCAGGTGACCGCATGGCTTGG 1301
| | | | |
OY 961 GACATGATGATGAATCTGTCCGCCACGCGCTACCATGATCTTGGGCTACGCGATGCTGTC 1020
| | | | |
DB 1302 GATATGATGATGAATCTGTGCTGCTCAACAGCCCTAGTGTGTGCAATTTACTCCGATC 1361
| | | | |
OY 1021 CCGGAGCTCATTAATGACATCATTTAGCGGGCTCATTTGGGGCTGTCATTTGGGCTTGGCC 1080
| | | | |
DB 1362 CCACAGCTGTCTGTAATGTGTGGCGGGGCCCTGAGAGTCTAGCGGGCTTGGCC 1421
| | | | |
OY 1081 TACTTCTCTATGACGAGGACGTGTGGCGAAGTGTGTCTCATCTCTTGTGGCGCGCGG 1140
| | | | |
DB 1422 TACTATTTCCATGTGTGGGGAATCTGGGCTTAAGTCTTGAATGTGATTTACTCTTTGGCGGC 1481
| | | | |
OY 1141 GTGAGCGCGGCAACCATATCTGTGTGGGGTCTGTGCCGCGAGACCAACCGGCGCTCAAC 1200
| | | | |
DB 1482 GTTGAAGGGGTCTACTTACACGACAGGGGGGTGCGAAGCTCCGACACACARGGGTCTGCG 1541
| | | | |
OY 1201 AGCTATTGATGATGGGCGCCGAGGAAATCCAGCTGTTTAAACCAATGGGACGCTGG 1260
| | | | |
DB 1542 TCTTCTTTCAACCHAGGGCGGCGCCAGAAATCCAGCTCATTAACACCAATGRCAGCTGG 1601
| | | | |
OY 1261 CACATCAACCGGACCGGCTGTAACTGCAATGACTCTTGGCACACCGGCTTTATCGCTCT 1320
| | | | |
DB 1602 CACATCAACAGGACTGGCTGTAACTGCRATGATCCCTCAACCGGGTCTTTGGCGGCA 1661
| | | | |
OY 1321 CTGTTCTTACACCAACAGCTTCAACTGCTGTGAGATGTCCGAAACGATGTCCGCTGCGGC 1380
| | | | |
DB 1662 CTGTTCTTACGCAACAAGTTCAACTGTCTCGGATCCCAAGCGCATGGCCAGCTGAGY 1721
| | | | |
OY 1381 AGTATCGAGGCTCTCCGGGTGGGATGGGGCGCTTGCAATATGAGGATTAATGTCACAAT 1440
| | | | |
DB 1722 TCCATATGACAGTGTGCTCAGGGAATGGGATCCCATCACTTAAACCGAG-----CTAAG 1775
| | | | |
OY 1441 CCAGAGATATGAGACCTTATTTGTGCACTTACCCACCAAGGACATGTGGCGTGTCTCC 1500
| | | | |
DB 1776 GACTGTGATCAGAGGCTTACTGCTGGCATTTAGGACCTCGGCACTGTGTATGCTACT 1835
| | | | |
OY 1501 GCGAAGACTGTGTGGCCAGTGTACTGTTTACCCCGGCGCATGTGTGTGGGACG 1560
| | | | |
DB 1836 GGGTGGCAGGTGTGTGTCTCAGTGTACTGTTTACCCCAAGCCGTGTGTGTGGGACG 1895
| | | | |
OY 1561 ACCGACAGGCTTGGAGCGGCGCACTTACAGTGGGGGGGAATGAGACAGATGTCTTCTTA 1620
| | | | |
DB 1896 ACCGATCGGTCCGGTCCCTTACGTATACGTGGGGGGCAATGAGACAGATGTGCTGCTT 1955
| | | | |

OY 1621 TTGAACAGACTGACCAACCGGTGGGATCATGTGCTGAGACGTGATGAACTTCTT 1680
| | | | |
DB 1956 CTGAACAAACCGGGCGCGGCAACAGGCAACTGGTTTGGCTGTACATGATGAAACAGCACT 2015
| | | | |
OY 1681 GGGTCAACCAAGACTTGGCGGCGCAACCCCTGCGTATCTAGAGCTGACTTCAACGCGAC 1740
| | | | |
DB 2016 GGGTTCACCAAGACGTGTGGGGGCGCTCCGATGCAATCGGGGGGTGGGCAACCTCA-- 2073
| | | | |
OY 1741 ACCGACTGTGTGGCCCAACGAGCTGTTTAAAGAACTCTGTATCTCACTTAACTCAAA 1800
| | | | |
DB 2074 -----CTTGAACCTGCCCAAGATTTGCTTTGGAGACACCCGAGGCACTTACAAAAA 2129
| | | | |
OY 1801 TGGGCTCTGGGCGCTGGCTCAGCGCAAGGTGCTGATGCACTACCTTACAGGCTCTG 1860
| | | | |
DB 2130 TGTGCTCGGGGCGCATGTGTTAACCTAGGTGTATGTGACTACCATACAGGCTCTGG 2189
| | | | |
OY 1861 CATTACCCCTGCACACTTAACTTACCATCTTCAAAATTAAGATGTATGTGGAGGGGTT 1920
| | | | |
DB 2190 CACTTACCCCTGCACTGTCACTTACATCTTTAAAGTTAGATGTATGTGGGGGCGTG 2249
| | | | |
OY 1921 GAGCAGGCTCAGGCTGATGCAATTTCACTGTGGGGATCGTTGCAACTTGAAGAGC 1980
| | | | |
DB 2250 GAGCAGGCTTATGTGCTGTGATGCACTGCACTCGAGAGAGGCTTGCACTGAGAGC 2309
| | | | |
OY 1981 AGAGACAAAGTCAACTGTCTCTTGTGTGCACTGCAACAGGAATGGGCAATTTTACT 2040
| | | | |
DB 2310 AGGACAGATCCGAGCTCAGCGCGCTGTGCTGTCCACGACAGATGGGCAACTTCCC 2369
| | | | |
OY 2041 TGTCTTTATCTGGAACCTGCGCGCTGTGTGACGTGTCTTCTGACCTCCACCAAACTTC 2100
| | | | |
DB 2370 TGTTCCTTCAACCACTTACCGGCTGTGTGCACTGTTGATTCACCTCATGAAACATC 2429
| | | | |
OY 2101 GTGAGAGTAAATTAATGATGAGCTATCACTGAGCTCACAATAATACATCGTCCGATG 2160
| | | | |
DB 2430 GTGAGCTGCAATCTGTACGTATAGGCTGACGCTGTCTCTTGTGTGATCAAAATGG 2489
| | | | |
OY 2161 GAGTGGGTAATTACTCTTATTTCTGTCTTGTAGCGAGCGCAAGGTTTGGCTGTATAG 2220
| | | | |
DB 2490 GAGTATATGATGTGCTTCTTCTCTGTGCGAGCGCGCGCTGTGCTGTGCTTGTGG 2549
| | | | |
OY 2221 ATGTCTATCTTGTGGGCGGAGCGGAAAGAGCACTGAGAAAGCTGTGCAATCTTGGACGCT 2280
| | | | |
DB 2550 ATGATGTGTGTATACCCAGGCTGAGGCGGCTGTAGAAACCTGTGTGTCTTCAACGCG 2609
| | | | |
OY 2281 GCGAGCGAGCTAGCTGCAATGAGCTTCTATATTTTGTCAATCTTTTGTGTGCTGTGG 2340
| | | | |
DB 2610 GCGTCSGTGCGGAGCGAGATGSCATTTCTCTCTCTGTGTGTGTGTGTGTGTGTGTGT 2669
| | | | |
OY 2341 TACATCAAGGCTGGGTAGTCCCTTGTAGCTATTCCTCACTGAGCGCTGTGTCTTT 2400
| | | | |
DB 2670 TACATCAAGGCGAGGCTGTGCTCTGGGGCGGCATATGCTGTGACCGGCTGTGGCGCTG 2729
| | | | |
OY 2401 AGCTTACTGTCTTACGATTTGCCCAAGGCTTATGTATGAGAGCATCTGTGCAATGGC 2460
| | | | |
DB 2730 CTCTGTCTCTGTGGGTGTACCAACAGGCTTATGCGCATGAGACGAGAGATGCTGCA 2789
| | | | |
OY 2461 CAGATAGGAGCGGCTGTGCTGTATGATCACTCTCTTTACTCAACCCCGGATTAAG 2520
| | | | |
DB 2790 TGTGTGTGAGAGGAGGTTTTTGTAGGTCTAATCTTGTGACTTGTGTACCACTACAA 2849
| | | | |
OY 2521 ACCCTTCTAGCGGTTTTTGT 2580
| | | | |
DB 2850 GTGTCTCTCGCAGGCTCATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2909
| | | | |
OY 2581 GTCAAGAGTGGGACCACTATGACAGTGGCGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2640
| | | | |
DB 2910 TGT 2969
| | | | |
OY 2641 GTGACCATATTTCAACCAAGGT 2700
| | | | |
DB 2970 ACATGTGCGGCGCAACCAAGCTAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3029
| | | | |
OY 2701 GGGCTGTGCTTACTCTTAAAGGTGCTTTGACGCGGCTGTGCTTGTGCTCAAGGCTCAC 2760
| | | | |

Db 3030 GGTCCGCTCAGTGGCTCCAGGCTGCCATACCTGCATGCGCTACTTTTGAACGGCTCAG 3089
Qy 2761 GCTTACTGAGATGTGACCATATGCAAGCATCTCCGGGGGAGTACCTCAAGT 2820
Db 3090 GGGTTATTTGTGATGATGATGTTGGTCCGAAAGTACCCGAGGCCATTAGCTCAATG 3149
Qy 2821 GCGCTACTAGCCCTTGGSCAGGTGACCTGACCTTATGATGACCACTTCAACCCCATG 2880
Db 3150 GCCTTATGAAAGCTGGCCGACCTGACAGATACGATCTATGACATCTTAACTCCGCTA 3209
Qy 2881 TCGATTTGGCTGTGATGAGCTGCGGAGCTTGGCGGTGCGCTTGAAGCTTATCATCTTC 2940
Db 3210 CAGGATTTGGGCCACCGGGGCTTACGAGACTTGGCGGTGAGTAAAGCCGCTGCTTC 3269
Qy 2941 AGTCCCATGAGAAAGAAAGTATTTGTGTGGGAGCGGAGACAGTCTGTTGGGACATT 3000
Db 3270 TCTGACATGAGACCAAGATCATCACTTGGGGGCGGACACCGCGGTGTGGGACATC 3329
Qy 3001 TTAACGGAATTCCCGGTGTCGCGGACCTTGTGGAGGTCCTCTTGGCCACGCTGAT 3060
Db 3330 ATCTTGGAGATCTCCCGTCTCCGCGCAAGGGAGATCTTCTTGGGACCGGCGGAT 3389
Qy 3061 GGGTATACCTCAAGGGGTGAGTCTTCTGCCCCATCACTGCTTACGCCACAGACACA 3120
Db 3390 ARCATTGAAGGAGCGGGGTGCGCATCTTGGCCCATCAAGGCTTACGCCCAACAGRCA 3449
Qy 3121 CGTGGCTTTTGGGACCATATGATGTGATGATGACAGGGGCGGACAAAGACAGAGCT 3180
Db 3450 CGGGGCTTACTTGTGTGATTTGTTACTAGCTTCAAGGCGGGACAAAGACAGGTGAG 3509
Qy 3181 GGGGAAATTCAGGCTCTGTCCACAGTCACTCAGTCCCTTCTCGGAACATCATCTGGGG 3240
Db 3510 GGGGAGGTTAAGTGGTTTCCACCGCAACAACTTCTTGGGACCTCGTATATGATG 3569
Qy 3241 GTTTTGTGACTGTCTACATGAGCTGAGCAACAAGACTTGGCGGCTCACGGGGTCCG 3300
Db 3570 GTGTGTGAGCTGTCTTCCACGGGGCGGCTCRAAGACCTTACGGGCCCAAAAGGCCA 3529
Qy 3301 GTACGCAATGTATCTCAAGTCTGAGGGGACTTATGATGGGTGGCCAGCCCCCTGG 3360
Db 3630 ATACCCAAATGTATCAATATGTAGACCAAGACTCTCGCGCTGAGTGGCTCCCGGG 3689
Qy 3361 ACTAAATCTTTGAGCGCTGTGACAGTGTGAGCGGTGACCTGTACCTGATCAGCGGAC 3420
Db 3690 GCGCTTCCCTTACACATGACCTGAGTACTCGGACTTATATGTGTACAGAGCAT 3749
Qy 3421 GCTGATGTCAATCCCGCTCGAAGACGGCGGACAAACGGGAGCGCTACTCTCCCGAGA 3480
Db 3750 GCTGATGTCAATCCCGGCGCGGACGGGAGCGGTGCGGGGAGCTACTGCTCCCGAGA 3809
Qy 3481 CCTCTTTCACCTTGAAGGGGCTCTCAAGAGCGCGGTGCTTATGCCCAAGGGGCACTG 3540
Db 3810 CCGGTCTCTTATTTGAAGGGCTCTTGGGTGGCCCACTGCTTGGCTTGGGGGCACTG 3869
Qy 3541 GTGAGGTCTTCCGGGAGCTGTGTCTCGGGGCGTGTGATGATCAATGATTTTCATC 3800
Db 3870 GTAGGATCTTCCGGGCTGTGTGTGACCCGGGGGTTGCGAAGGGCGGTGACTTATTA 3929
Qy 3601 CCCGTTGAGACCTTCAACATGTCAGCGGCTCCCGCACTTTATGTAACAACACACACA 3660
Db 3930 CCGGTTGAGTCTATGAGAACTACTATRCGGTCCCGGCTTTCACGGAACATCATCTCC 3989
Qy 3661 CCTGCTGTGCCCAAGCATATCAGGTGCGGATCTTGATGCCCGGACCTGGAGTGAAG 3720
Db 3990 CCGGCGGTACCGGAGCATTTCAAGTGGCCATCTACCGCTCCCACTGGCACCGGCAAR 4049
Qy 3721 AGCACAAAGTCTCTGTGATATGCTCAGGGGATATAAGTGTAGTAACTTAACTCC 3780
Db 4050 AGGACATTAAGTGGCGGTGCTGAGAGCCCAAGGGTACAGGATCTGCTCTGMAACCA 4109
Qy 3781 TCACTGCTGCAACCTTGGGGTTTGGGCGTACTTGTATAGGACATGGCATCAATCC 3840

Db 4110 TCTGTGCCGCACTTTRGTTTTGGGGCGTATATGTCTAAGGACATGGTACGACCCC 4169
Qy 3841 AACATTAGACTGAGTCAAGACTGTGACGACCGGGGCGCCCTACAGTACTCAATAT 3900
Db 4170 AACATCAGACTGGGGTAAAGAACATTAACACGGGGCGCCCACTACGTACTCACCTAT 4229
Qy 3901 GGCAAATTCCTCGCGATGGGGGCTGTGCGGGGCGGCGCCCTACAGCATCATATATGAT 3960
Db 4230 GTTAACTTCTTCCGACGCTGATGATTTCTGGGGGCGCTTATATATCATATATGAT 4289
Qy 3961 GAATGCCATCCGCTGACCTTACACCATCTTGGATCGGAACAGTCTTGTATCAAGCA 4020
Db 4290 GAGTGCACTCACTGACTGATCACTTCTGGGATGGGACAGTCTCGACCAAGCG 4349
Qy 4021 GAGACAGTGGGGTCAAGCTTAACTGTGTGCTGACAGCTACGCCCCCTGGGTCAGTACA 4080
Db 4350 GAGACGCTGAGACCGGGCTGTGCTGTGCTGACACCGCTACGCTCGGATCGTAAAC 4409
Qy 4081 ACCCCCAACCCCAACATAGAGAGTGGCCCTTGGGAGAGGGCGGATCCCTCTAT 4140
Db 4410 GTGCCACACCCAAATTTGAGAGTGGCTCTGTCCAACTGAGAGATCCCTTCTAT 4469
Qy 4141 GGGAGGGCAATTCCTGTCTTACATCAAGGAGAAAGATCTGATCTTCTGCAATTA 4200
Db 4470 GGCAGGGCAATCCCTCTCGAGGCTACAGGGGGGAAAGGATCTATTTTCTGCACTCC 4529
Qy 4201 AAAAAAAGTGTACAGAGCTCGCGGCGCCCTTGGGGTATGGGCTTGAATCAAGTGGCA 4260
Db 4530 AAAAAAAGTGTACAGAGCTCGCGGCGCCCTTGGGGTATGGGCTTGAATCAAGTGGCA 4589
Qy 4261 TACTACAGAGGTTGAGAGCTCTCGATATCAAACTCAGGAGACGATGAGTGGTCCG 4320
Db 4590 TATTTACGGGGTTTGAAGTGTCTGTCTGATACCACTGAGGAGATGCTTATGTGGCA 4649
Qy 4321 ACCGACGCCCTCATGACAGGGTATATCTGGGAACTTTGACTCTCGATGCACTGCAACGTA 4380
Db 4650 ACAGACGCTTATATGACGGGCTTACACGGGATCTTGAATCGGATGATGCACTGCAACA 4709
Qy 4381 GCGGTCACTCAAGTGTATGACTTCAAGTATGACCCCACTTCACTCAACACAGATT 4440
Db 4710 TGTGTACCCAGACRGTGTATTCAGCTTGAATCCCACTTTACATGAGACGACAC 4769
Qy 4441 GTCCCTCAAGAGCTGTCTCAGTACGACGCGCGGGGTGCAAGGGTGGGGAGACTG 4500
Db 4770 GTGCCCAAGACGGGTTGTGCTCAGTACAAAGGAGGAGGATGCTGAGGGTATGGGA 4829
Qy 4501 GGCATTTATAGTATGTTTCACTGAGTACGAGCTCTAGATGTTTGAAGTGTATG 4560
Db 4830 GGCATTTACAGGTTTGTGACTCGGAGAAACGGCCCTCGGGCATGTTGATCTTCTG 4889
Qy 4561 CTCTGTGAGTGTACGACGAGGGGCGCATGTATGAGCTCACACATCGAGACAC 4620
Db 4890 CTGTGTGAGTGTATGACCGGGCTGTGTGTGTATGAGACTCACGCGCTGAGACTTACA 4949
Qy 4621 GTGAGGCTCAGGGCGTATTTCAACAGACCGGGTTGCTGTGTGCGCAAGCACTTTGAG 4680
Db 4950 GTGAGGTTGCGGCTTACCTGATATACCRGGGTTGCCGTCTGCGAGGACATCTGAG 5009
Qy 4681 TTTTGGAGGAGTTTTCACCGGCTCTACACATATAGTCCCACTTCTTCCAAAGA 4740
Db 5010 TTTCTGGAAAGGTTTTCACAGGCTTACACCACTATAGTCCCACTTCTGTTCCAAAC 5069
Qy 4741 AAGCAATGGGGGAAATTTGCGATCTTAAAGCTTACAGGCTACAGGCTACAGTGTGCTAG 4800
Db 5070 AAGCAGGACAGAGAACATTTCTTACCTGTAGCTTACCAAGCCACGGGTGTGCGCAG 5129
Qy 4801 GCCAAAGCCCCCGCTGCTGAGGAGCTATGTGGAAGTGTGCTGCACTCAAGGCC 4860
Db 5130 GCTCAGGCTTCACTCTCATGCTGTGAGTCAAAATGTGAAATGTCTATACGCTTAACT 5189
Qy 4861 ACACTGTGGGCCCAACACTCTCTGTATCCGCTTGTATGAGGCTGTATCAACAGAGTCA 4920
Db 5190 ACGCTGCAAGGGACCAACCCCTGTGTATGAGGCTGTGAGGCGTTCAAATATGATCAAC 5249

OY	4921	TTCAACAATCCGGACGAATAATATGCAACCTGCAATGCAAGGCGACCTTAGATCATG	4980
Db	5250	CTCAACAATCCCAATACCAAAATTCATCAATGGCAATGATGTGGCCGAATCGAAGATGTCTC	5309
OY	4981	ACCAGACATGGGCTTTGGCAGGGGAGATCTTTGGCGGCGCTGCGCGCTATTTGGCTTGAGC	5040
Db	5310	ACTAGCACTCTGGGTGCTGTAGGGGAGATCTTGAGAGCTCTGGCCGCAATATTTGCTGACA	5369
OY	5041	ACCGGATGTGTTTGCATATGAGCCGCTTGGCAATTTAACGCGAGCCGTGTTTGGCCCG	5100
Db	5370	ACCGGATGTGATCATATTGTGGGTAGATTCATTTTGTCCGGAGAGCCGGCTGTTGTTCCC	5429
OY	5101	GACAAGAGGTCCTCTATGAGGCTTTTATGATGAGGAATGAGCTTCAAGGCGCGCT	5160
Db	5430	GACAGGAAGTCTCTTACCGGAGATTGCAATAAATGAAAGATGCGCTTCACACTCCCT	5489
OY	5161	CTCATTTGAAGGGGACAGCGGATAGCCGAGATGCTGAATGCCAAGTCCAAAGCTTATTTG	5220
Db	5490	TACATCGAACAAGGAATGACAGTCCGCGAGAGTTCAAGCAGAAAGCACTCGGGTTTGCTG	5549
OY	5221	CAGCAAGCTTCCAAACAAAGCTCAGACATACAAACCACGTGGCAGGCTTCAATGGCCCAAG	5280
Db	5550	CAACAGCAACCAAGCAAGCGAGGCGCTGCTCCCGTGGTGGATCCAGGTGCGGGCC	5609
OY	5281	GTAGAACAATTTCTGGGCGCAACACATGATGGAATCTTATTAAGGGATCCAAATACCTGCA	5340
Db	5610	CTTAGAGCCTTCTGGGCAAGACCAATGTGGAATCTTCAACAGGGGATTCAGATATTTAGCA	5669
OY	5341	GGACTATCAACACTGCGCAAGGAAACCTTCAGTAGGCTTCATGATGGCGTTGACGTGCGCG	5400
Db	5670	GGCTTATCACTCTGCTGTGGGACC CGCGATAGATCATGATGGAATTTCAAGCCTCT	5729
OY	5401	CTCACCAATCCGCTGTCAACAAGCAACCATCTCTTCTCAACATTTTGGGGGCTGTGCTA	5460
Db	5730	ATCAACAATCCGCTCACCAACCCAGAAATCCCTCCTTTCAAATCTTATAGGGGAATGGGTG	5789
OY	5461	GCATCCCAAAATTGACCAACCGCGGGGGGCACTGGCTTGTTGTCAGTGGCTATGTTGGA	5520
Db	5790	GCTGCTCAACTGCTCTTCCCAATGTCCTCTTGCRCTTGTGTGGGTGCCGCAATTTGCCGAT	5849
OY	5521	GCTGCGGTAGGAGATATAGAGGCTTTAGTAAAGTGTCTAGTGAACATCTTGGCAGGATAGT	5580
Db	5850	GGCGCATTTGGAGATAGGCTTGTGGAAAGTGCTGTGTGACATTTCTGGCGGGCTATGGA	5909
OY	5581	GGGGGCATTTGGGGGGCTCTGTGCCATTTCAAATCATGTCTGGCGAGAGCCTTCATG	5640
Db	5910	GGGGGGGTGGCGGGTGCACTGTTGGCTTTTAAAGTCATGAGCGGGGAGACCCCTCCGCC	5969
OY	5641	GAGATGTGTCCTCAATTTGCTGTGCTGGAATTTCTGTCCGGGGTGCTTGTATGTGGAATC	5700
Db	5970	GAGAACCTGGTTAATCTTGCTTCTGTCGCATCTCTCCCGGGGCTTGTATGTATGATGATC	6029
OY	5701	ATCTGCGGGGCCATTTCTGCGCGGACACAGTGGGACCGGGGGAAGCGCGCTCCAAATGGAATG	5760
Db	6030	GTGTGTGAGCAATCTGTGGTGGGACGATGGGCCCGGGAGAGGGGGCTGTGCAATGGAATG	6089
OY	5761	AATAGACTATTTGCTTGTCTTCCAGAGAAATCAAGTGGCCGCCCACTAGATGACG	5820
Db	6090	AACCGCTGATAGCCTTGGCTTCCGGGGGTAAACAAGTTTCCCCCAACCACTACGTGCT	6149
OY	5821	GAGTGGATGCGTCCGAGGGTGTAGCCCAACTACTTTGGCTCCCTTAACATTAACAGCCTG	5880
Db	6150	GAGAGCAAGCGCGCAGACATGTAACTAGATCTCTCTGAGCTTCAACATCACTGACGTTG	6209
OY	5881	CTCAGAAACTTCCACAATGGAATTACTAGAGACTGCCCCATCCATGCGGGGCTGTG	5940
Db	6210	CTGAAGAAGCTTCAACAATGGAATTAATAGGACTGTCTCAAGCAATGCTCGGGCTCGTGG	6269
OY	5941	CTCCGCAATGTGGGACTGGGTTTGCACATCTTAACAGACTTTAAAAATTTGCTGACC	6000
Db	6270	CTAAAGGATGTTTGGGACTGGATATAGCAAGGTGTTGACTTGACTTCAMAGCCTGCTCCAG	6329

QY	6001	TCGAAATATTTCCGAAAGATGCGCGGCTCCCTTTGTCCTGTCGAAAGGGGTACAG	6060
Db	6330	TCGAGCTCTCGCGGAATTGCGGGAGTCCCTTTCTTCATGCGCAAGCGGGTCAAG	6389
QY	6061	GCGCTGTGCGCGGCACTGCGCATATACCAACGAGTGTCTTTGCGGCGCCATATCTCT	6120
Db	6390	GGAATCTGTGCGGGGGAGCGGATCATCGAAACCACTGTGCCATGTGAGGACAAATTACC	6449
QY	6121	GGCAATGTCCGCTTGGGCTCCATGAAATACCGGGGCTTAAAGACTGTGATATCTGG	6180
Db	6450	GGAATGTCAAAAACGGTTCCATGAGATCGTTGGGGCTTAAACCTTCAGACACCTGG	6509
QY	6181	CAGGGGACCTTTCCATCAATTTGTTATCACGGAAGGGCCAGTGCCTGCCGAAACCCGCGCA	6240
Db	6510	CACGGAACGTTTCCCATTCACCGGTACACACAGGCCCTTCACACCTCCCGGCGCG	6569
QY	6241	AACTTAAAGTGCGCATCTGGAGGGTGGCGGCTCAGAGTACCGGAGGTGACCGACAC	6300
Db	6570	AACTATTCCAGGGGCGTGTGGCGGGTGGCTGTGAAGATACGTGAGATTAGCGGGTG	6629
QY	6301	GGGTTCATTCACATACATACAGACTACCACTGATTAATTGAAAGTCCCTGCCACTA	6360
Db	6630	GGGGAATCTTCCACTACGTGACGGGTATGACCAACGACAAAGTAAATCCCGTGCAGATC	6689
QY	6361	CCCTCTCCCGAGTCTTTTCTGGGTGAGCGAATGACATCCATAGTTTGGCCCCACA	6420
Db	6690	CCGGCCCCGAAATCTTTTACCGAAATGGAACGGGTGGGTTGACAGATACGCTCCGGCG	6749
QY	6421	CCGAAGCCGTTTTTCCGGATGAGAGTCTCGTTCCTGGCTTAAATTCTTTGTGCTC	6480
Db	6750	TGCAGACTTCTTCACCGGAGAGTGCACATTCAGAGTCGGGCTCAACCAATACCTGGT	6809
QY	6481	GGGTCCAGCTTCTTGGGACCTTGAACCCGACACAGACGTATTGATGTCCATGCTAACA	6540
Db	6810	GGGTACAGCTCCCATGCGAGCTCGAACCGGATGTGCAAGTGCCTCACTTCATCTCAC	6869
QY	6541	GATCATCTTCATACACGGCGGAACTGTGACGGCGGCTTAAAGGCGGGGTACCCCCA	6600
Db	6870	GACCCCTCCCAATTACAGCAGAAACGGCTAAAGTATGGCTGGCCAGGGGGTCTCCCC	6929
QY	6601	TCCGAGGCAAGCTCTCGGCGAGCGAGCTATCGGCACCATCGCTGCGAGCCACTGCACC	6660
Db	6930	TCTTTGGCAGCTTTCAAGTACGCAATTGTCTGGCTTCTCTTGAAGGCAATGCACCT	6989
QY	6661	ACCACGCGCAAGCCTATGATGTGACATGCTGTGATGTCACT-----GTTC	6708
Db	6990	ACCACCATGACTCCCCGAGCGTGCACCTATCGAGGCGCACTCTGTGCGGACAGAG	7049
QY	6709	ATGGGGGGCGATGTGACTCTGGATAGAGTCTGTGGTCCAAAGTGTCTGTGGACTCTCTC	6768
Db	7050	ATGGCGGGAACATCACCCGTGTGGAGTCAAGAAATAGAGTGATTTTGGACTCTTTTC	7109
QY	6769	GACCCAAATGATGGAAGAAAGGAGCGACTTAGCTTCGATACATCAGATTAATCATGCTC	6828
Db	7110	GAACCTCTTCAAGCGGAGAGAGATGAGGGGAAGTATCTGTTGACGAGAGATCTCGGA	7169
QY	6829	CCCAAGAAAGAGTTTCCACACAGCTTTACCGGCTGTGGACCGGCTGATTAACAACCCACG	6888
Db	7170	AAATCCAAAGAGTTTCCCCCGCGGTGCGCATATGGGACCGCCGGAATTAACAACCTCCA	7229
QY	6889	CTTGTGGAATCTGTGAAAAAGCCAGATTACCAACCGGCACTGTGTGGGGCTGTGTCTC	6948
Db	7230	CTGTTAGAGTCTGTGAAAGTCCGGAATCACTCCCTCCGGGCGTGTGCAATGGTGTGCCATTG	7289
QY	6949	CCTCTCTTAGAAAAACCCCGACGCTTCCCAAGAGAGCGCCGGAACGTGTGGCTTAAGT	7008
Db	7290	CCGCTTACACGGGCGCTTCCAAATACCGCTTCAAGAAAAAGAGACGTTTGTCTGACA	7349
QY	7009	GAGACTCATAGAGAGATGACCTTTCACAGCTGTGCATTAAATGCTTTTGGCCAGCCCCC	7068
Db	7350	GAGTCCACCGTGTCTTCTGTGCTTGGCGGAGCTGTGCTACTAAGACATTTTCCGAGCTC---	7405
QY	7069	CCAAACGGCGAATTCAGGCTTTTCCAGGGGCGGGCGCTGCTCGATTCCGCGAGTACAGC	7128

Db 7406 -----CGGATGCTGGCCGCTTGACAGGGGACCGGCCACCGCCCTCCCGA---TCAGACC 7457
Qy 7129 CTTCTGATGATGTTGGCCCTTTTGGAGACAGGTTCCATCTTTCATGATGCCCCCTTCGAG 7188
Db 7458 TCTGACAGACGGGTGACAAAGATCTGACGTTGAGCTGACTCTCCATGCCCCCTTGAG 7517
Qy 7189 GGGGAGGTGGAGATCCAGACCTGGAGCTGAGAGAGGTAGAGCCCCCAACCCCCCCAG 7248
Db 7518 GGGGAGCCGGGGACCTTGATCT----- 7540
Qy 7249 GGGGGGTGGCAGCTCCCGGCTCGAGCTGGGGTCTGCTCTACTTTCCTCCGAGAGAC 7308
Db 7541 -----CAGCGAGGGGTCTTGCTTACCGTAGGGGGAGGCCAGC 7580
Qy 7309 GACTCCGCTGCTGCTGCTCTCATCTCACTCTGAGACGGGGGCTCTTAATTAATCTTGT 7368
Db 7581 GAGGACATCTGCTGCTCTCAATGCTCTACACATGAGCAGGGGCCCTGATCACGCCATGC 7640
Qy 7369 AGTCCCGAAGAGAGAGTTAACCGATTAACTCCCTTGAGCACTCCCTGTTGGCATATCAC 7428
Db 7641 GCGGACAGAGAGAGCAAGCTGCCCATCAAGCGTTGAGCACTCTGCTGCTGACACAC 7700
Qy 7429 AACAGGTGTACTGTACCAACAAGAGCGCTCACTAAGGGCTAAAGATTAATTTT 7488
Db 7701 AACATGCTCTATGCGACAAACATCTGAGAGCAAGCTTGGCGAGAGAGAGTCACTTT 7760
Qy 7489 GATAGATGCAAGTGTCTGACTCTCTACTACACTCACTCTTAAGGACATTAAAGCTTAGC 7548
Db 7761 GACGACATGCAAGTCTGAGACCACTACCGGGACRTGCTCMAGAGATGAAGGCCAG 7820
Qy 7549 GCTCCGAAGTCAACCGAAGCTCTCCATAGAGAGAGGCTTCCAGTTAAACCCACCC 7668
Db 7821 GCGTCAAGTAAAGGCTTAACTTCTATCTGTAAGAAAGCTCAAGCTGACGCCCA 7880
Qy 7609 CATTTGCAAGATTAATATATGAGTTTGGGCTAAGAGGTCCGACGTTTCCGGAGG 7668
Db 7881 CATTCCGCCCAATTAATTTGGCTATGGGCAAGAGACGTCCGGAACCTATCCAGAG 7940
Qy 7669 GCCGTAAACCATCAAGTCCGTGTGAGAGACCTCTCGAGAGACTCAAGAAACCAATT 7728
Db 7941 GGCATTAACCAATCCGCTCGGTGTGAGAGACCTGCTGAGAGACATCTAGACACAAAT 8000
Qy 7729 CCCCAACCATTAATGCGCAAAAAAGAGTCTTCCGTGACCCCAACAGGGGGGCAAG 7788
Db 8001 GACACCAACATCAAGGCAAAAAAGTAATTTCTGCTCAACCGAGAAAGAGGCGC 8060
Qy 7789 AAAGCAGCTGCGCTTATGTTTACCTGACCTCGGCGTCAAGGCTGCGAGAAATGACC 7848
Db 8061 AAGCAGCTGCGCTTATGTTTACCTGAGACCTGAGAGTTGCTGATGCGAGAAATGACT 8120
Qy 7849 CTTTATGACATTAACAACAACTTCTCAGGCGGTATGAGGAGCTTCTTATGATTCAG 7908
Db 8121 CTTTACAGACGTGCTCCACCTTCTCTCAGGCGGTATGAGGCTCTGTAACGATTCAG 8180
Qy 7909 TATTTCCCGCTGACGGGTATGAGTTTCTTGAAGAGATGGGGGAAAGAAAGACCT 7968
Db 8181 TACTCTCTGGGCGAGGGGTGAGTTCTGATGATGCTGGAATCAAGAAAGAGCCCT 8240
Qy 7969 ATGGGTTTTTGTATGATACCGAGATGTTGATCTCAACCGTCACTGAGAGATGAG 8028
Db 8241 ATGGGCTTCTGTATGACACCGGCTGCTTGAATCTCAACAGTCACTGAGAGATGCT 8300
Qy 8029 ACTGAGAGTCCATATATCGGGGCTGCTCTTGGCCGAGAGGCCCACTGCAATAC 8088
Db 8301 GTTAGAGAGTCAATCTACATGTTGATCTTGAGCTGCCCCCGAGACAGAGCATTAAG 8360
Qy 8089 TCGCTAATGATGAGACTTTATGAGGAGGCGCTATGTTTCAACAGCAAGGCGCAACCTGC 8148
Db 8361 TCGCTTACGAGGCGGCTTATCATCGGGGCGCCCTGACTAATTAAGAGGCGCAATGCTC 8420
Qy 8149 GGGTACAGGCGGTTGCGCGCCAGCGGGGTGCTCACTAGCATGAGGAGAACCATCA 8208

Db 8421 GGTTATCGCCGCTGCGCGGAGGCGGTGCTGACGACTAGTGGGCAATACCTCACA 8480
Qy 8209 TGTACGTGAAGCCTTAGCGGCTTGTAAAGCTGCAAGGATTAATCGGCCCAATCTG 8288
Db 8481 TGTCTATTGAAACCACTGGGCGCTGTGAGCTGCAAGGCTCCAGACTGACATCTC 8540
Qy 8269 GTAAGCGGATGACTTGTGTATCTCAGAAAGCCAGGGGAGCCGAGAGAGAGAGCG 8328
Db 8541 GTGAGCGGAGACGACTTGTGTATCTGTGAAAGCGGGAGAACCGAGAGATGCGCG 8600
Qy 8329 AACCTGAGACCTTACAGAGGCTATGACCAAGTATTCCTCTCTGTGACCCCC 8388
Db 8601 AGCTTACGAGCTTACAGAGGCTATGACTAGGTACTGCTCCCGCCCGGGAGACCCGCC 8660
Qy 8389 AGACCGAGATATATCTGAGCTGATTAACATTTGCTCTCTCAATGCTGTGGGCTG 8448
Db 8661 CAACCAAGAAATGACCTTGGAGTGAATCAATGCTCTCAACGTTGCTGCGCAG 8720
Qy 8449 GGGCCCAAGGCGCCGACATATCTACTGACCAAGAGACCTTACACTCCATGCGCCG 8508
Db 8721 GAGCATTCCGCAAGAGAGTACTACTGATCTGATCCACACCCCTTGGACAG 8780
Qy 8509 GCTGCTGGGAAAACATTAAGACACTCCCTGTCAATTATGCTGGGAAAATCATCATG 8568
Db 8781 GCTGCGTGGAGACAGCTTACACACTCCAGTCAACTCTGGCTAGGCAATCATCATG 8840
Qy 8569 TAGGCCCCGACCATATGGGCTGCGATGCTCTGATGACACATCTTCTTCATCTCATG 8628
Db 8841 TAGGCRCCCACTTATGGGAGAGATGATTTGATGACTCACTTCTCTATCTCTCTA 8900
Qy 8629 GCTCAAGACAGCTGAGACCAAGACTCACTTTGAGATGATAGGAGCGGTACTCCGTG 8688
Db 8901 GCCCAGAGARAGCTTGAAGAGCCGAGATTTGTCAGATCTACGGGCGCTGTACTCAT 8960
Qy 8689 AGTCCCTTGGACCTCCGAGCTATTAATTAAGGTTAATGATGATGATGATGATGATG 8748
Db 8961 GAGCCTTGTACCTTACCTCAGATCATTAATCAACGACTCATGATGATGATGATGATG 9020
Qy 8749 CACATATCACTCCCAACGAACTGACAGCGGTGCTTACCCCTCAGAAACCTTGGGGCG 8808
Db 9021 CATAGTTACTCTCAGGTGATCAATAGGTTGCTTCAAGCTCAGAGAACTTGGGGTA 9080
Qy 8809 CCAACCTCTGAGGAGTGAAGAGCGCGGACAGTGCAGTCAAGGCGCTCATCTCCGT 8868
Db 9081 CCGCTTGTGKAGTCTGAACATCGGACCAAGTCTCGCGCTAGCTACTCTCCAR 9140
Qy 8869 GGGGAGAGAGCGCGGTTTGCAGTCAATCTCTTCAATTTGGCGGTGAAGACCAAGCTC 8928
Db 9141 GGGGAGAGGCGCGVAGTGTGGAAGTACTCTTCAACTGGGAGTAAAGGACCAARCTC 9200
Qy 8929 AAATCACTCATTTGCTGGAAGCGCGCTCTGATTTATCAAGTGTTCACCGTGGC 8988
Db 9201 AAATCACTCAATCCCGGCGGCTCCAGCTGACATGTCRCGTGCTGTTGCTGT 9260
Qy 8989 GCGGCGGGGCGCAATTTATCAAGGCTGCGGTGCGGACCCCGCTTATGCTCTT 9048
Db 9261 TACAGCGGGAGACATATATCACAGCTGTCTGTCGCCGACCCCTGTGTTATGCTG 9320
Qy 9049 GGCCTACTCTTCTTGTGAGGGTGAAGGCTTTTCTACTCCCGCTCGG 9099
Db 9321 TGCTTACTCTTCTTCCGTAGGGGTAGGATCACTGCTCCCAACCGG 9371

RESULT 12
AAD2517
ID AAD2517 standard; DNA; 9413 BP.
XX
XX AAD2517:
XX AC
XX 26-MAR-2002 (first entry)
DT
XX
XX Hepatitis C virus complete DNA genome.
XX

KM HCV; hepatitis C virus; cytostatic; cancer; immunosuppressive; virucide;
KM antibacterial; fungicide; protozoicide; antipneumatic; antiinflammatory;
KM antithrictic; rheumatoid arthritis; neuroprotective; multiple sclerosis;
KM immune response; vasotropic; vaccine; gene therapy; autoimmune disease;
KM vasculitis; ds.
XX
OS Hepatitis C virus.
XX
PN MO200176643-A1.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-US011372.
XX
PR 07-APR-2000; 2000US-0195680P.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX
PI Orson FM, Kinsey BM, Bhogal BS;
XX
DR WPI; 2002-066308/09.
XX
PT Composition for oral delivery of vaccines, comprises expression vector
PT containing antigenic genomic sequence, bound to aggregated protein-
PT polyclonal polymer conjugate or suspension.
XX
PS Disclosure; Page 84-87; 145pp; English.
XX
CC The invention relates to a composition comprising an expression vector
CC bound to an aggregated protein-polyclonal polymer conjugate or
CC suspension. The expression vector contains a promoter polynucleotide
CC sequence operatively linked to a polynucleotide sequence encoding an
CC antigen which is a fragment of a gene or genome associated with an
CC infectious disease, cancer and autoimmune disease such as rheumatoid
CC arthritis, vasculitis, and multiple sclerosis, pathogenic genomes
CC consisting of bacterium, fungus, protozoa and virus such as human
CC immunodeficiency virus (HIV), herpes simplex virus (HSV), hepatitis C
CC virus (HCV), influenza and respiratory syncytial virus (RSV), and
CC optionally comprising a nucleotide sequence encoding a cytokine (or a
CC cytokine expression vector), is useful for inducing an immune response
CC (systemic and/or mucosal) in an organism. The cytokine expression vector
CC contains a sequence for granulocyte macrophage-colony stimulating factor
CC (GM-CSF) or interleukin-12 (IL-12). The polynucleotide sequences encoding
CC the antigen and the cytokine are under transcriptional control of same or
CC different promoter polynucleotide sequences. The expression vector, as a
CC DNA vaccine is useful for treating a condition in an organism. The
CC present sequence is hepatitis C virus complete DNA genome related to the
CC invention
XX
SO Sequence 9413 BP, 1886 A; 2816 C; 2694 G; 2017 T; 0 U; 0 Other;
Query Match 45.5%; Score 4141.8; DB 6; Length 9413;
Best Local Similarity 66.6%; Pred. No. 0;
Matches 6066; Conservative 0; Mismatches 2952; Indels 93; Gaps 6;
QY 1 ATGAGCAAAATCTTAACTTAAAGAAAACCAAAAGAAACACCAACCTGCGCCACAA 60
DB ATGAGCAAAATCTTAACTTAAAGAAAACCAAAAGAAACACCAACCTGCGCCACAG 389
QY GACGTAAAGTTTCGGGGGGGGGGGCGAGATCGTTGGCGGAGTACTTGTTCGGCGGAGG 120
DB GACGTAAAGTTTCGGGGGGGGGGGCGAGATCGTTGGAGTTTAACTGTTCGGCGGAGG 449
QY 121 GGCCCAAGTTGGGTGGCGCGCAAGAAAGACTTGGAGCGGTCCGACCAAGTGA 180
DB GGCCCAAGTTGGGTGGCGCGCAAGTGAAGAACTTCGAGCGGTGCAACTCTGTGA 509
QY 181 AGGCGCCAGCCCATCTTAAAGATGGCGCTCCACTGGCAATCTTGGGAAAAACAAGA 240
DB AGGCGCAACCTATCTCCAAAGGCTGCGCGCCGAGGGGTAGAGACTGGGCTCAAGCCGGG 569
QY 241 TACCCCTGGCCCTTATAGGGAAATGAGGAGCTCGGCTGGGAGATGGCTCTGTCGCC 300

DB 570 TACCTTGGCCCTCTATAGGCAACGAGGATATGGGGTGGCAGAGATGCTCTGTACACC 629
QY CGAGTTTCCCGTCCCTCTTGGGGCCCAATGACCCCGGCAATAGTGGCGCAACGTGGAT 360
DB CGTGGCTTCGGCTAGTGGGGGCCCAAGAACCCCGGCGTAAAGTGGGTAAATTTGGAT 689
QY 361 AAGTCATGATATCCCTTAAAGTGGGCTTGGGACCTTATGGGGTAACTCCCTGTCTG 420
DB AAGTCATGATATCCCTTAAAGTGGGCTTGGGACCTTATGGGGTAACTCCCTGTCTG 749
QY 421 GGCGCCCGGCTGGGGGGGGCGGAGAGCTTCGCGCATGGGCGTGAAGTCCGAGAGAC 480
DB GGCGCCCGGCTTGGGGGGGGCGGAGAGCTTCGCGCATGGGCGTTCGAGAGAC 809
QY 481 GGCGTTAATTTTGCACAGGAACTTACCGGTTGCTCTTTTCTATCTTCTTGTCAGCC 540
DB GGCGTTAATTTTGCACAGGAACTTACCGGTTGCTCTTTTCTATCTTCTTGTCAGCT 869
QY 541 CTGCTGTCTGATCACCAACCCCGGTCTCGCTGCCGAAGTGAAGAACTCAATACCGGC 600
DB TTGCTGTCTGTTTGAACATCCAGCTTCGCTTACGAGGTGGCAACGTTCGGGAT 929
QY 601 TACATGTGATTAAGCATGACGACCAATGACAGCATTAACCTGGAGCTCCAGGCTGCTG 660
DB TACATGTGATTAAGCATGACGACCTTCAACTCAAGTATTTGATGAGGACAGGACATGATC 989
QY 661 CTCACAGTCCCGGGTGGCTCCGTCGAGAAAGTGGGGAATGATCTGAGTGGATGA 720
DB ATGCAACACCCCGGGTGGCTCCGTCGAGAAAGTGAATTTTCCGTTGCTGGAT 1049
QY 721 CGGCTTACCGGAATGTGCGTGGACGGCCCGGCTTACGAGAGGCTTGGGAGC 780
DB GCGCTCATCCACGCTCGCGGCAAGAAACAGAGCATCCCAACGCAATACACAGC 1109
QY 781 CACATGACATGGTTGATGATGTCGCGACGCTGCTGCTGCTTACGTTGGGGAGCTC 840
DB CACATGATTTGCTGTTGGGGAGGCTGCTCTCTGTTCCGCTATGATGATTTGGGATCTC 1169
QY 841 TGGCGTGGGGTGAATGCTGCGACGCCCAATGTTTATTTGTCGCGGACGACCATGTGTT 900
DB TGGCGATCGTTTCTGCTGCTGCGGAGGCTGCTCTCTGTTCCGCTATGATGATTTGGGATCTC 1229
QY 901 GTCCAAAGCTGCAATGCTTCATCTTACCTGTGATCATATGAGAACCGCATGATGATG 960
DB GTCCAAAGATTTGCAATGCTTCATCTTACCTGTGATCATATGAGTACCGCATGATGCTTGG 1289
QY 961 GACATGATGATGAACTGTGCGCCACGAGTACCATGATCTGGCGTACGATGCTGTC 1020
DB GATATGATGATGAACTGTGCGCCACGAGTACCATGATCTGGCGTACGATGCTGCGATC 1349
QY 1021 CCCGAGGTCAATTAAGCATATTAAGCGGGGCTCATTTGGGGGCTCATGTTGCGTTGGCC 1080
DB CCACAAAGCGGTGATGATGAGGCGGGGCCCACTGGGGGTGCTTCAAGCGGGCTTGGCC 1409
QY 1081 TACTTCTTATGAGGAGCGTGGGGAAGTGTGTTGATCTTCTTGTGGCGCGGG 1140
DB TACTTATTCATGTTGGGGAATGAGGCTTAAGTCTTATGTTGATCTTGTGTTGCTG 1469
QY 1141 GTGAGCGCGGCAACCATCTGTTGGGGGTTTTCGCGCGAGACCAACCGGCGCTCACCC 1200
DB GTTGAAGGAGCACCACTGATGAGGAGGGAAGGTAGCTCCAGCACCCAGAGCTCTGTG 1529
QY 1201 AGCTTAATTTGACATGGGCCCCAGGCAAAATTCAGCTGTTAAACCAATGGCACTGG 1260
DB TCCTGGCTTCACAAAGGCCATGTCAGAAATTCAGTCTGTGAACCAACAGCACTGAG 1589
QY 1261 CACATCAACCGCACCGGCTGAACTGCAATGATCTTGTGACACCGGCTTTATCGGTCT 1320
DB CACATCAACAGGACCGGCTGAAATGCAATGATCTTGTGACCACTGAGTTCAATGTTGCG 1649
QY 1321 CTGTTTCAACCAAGCTTCAACTGTCAGATGTCGCGAAGCATGTCGGCTGCGCCG 1380
DB CTGTTTCAACCAAGCTTCAACGCTGCGGATGCTCCAGAGGCAATGGCTAGCTGCGCC 1709

QY	1381	AGTATCCAGAGCCTTCCGGGTGGGATATGGGGCGCTTGCATATATGAGATATATGTCAACCAAT	1440
Db	1710	CCCATTCGATGAGTTGCTCAGAGGGGTGGGGTCCCATCACTCATATGATATGCTCAGAGCTC-	1768
QY	1441	CCAGAGATATAGAACCTTATTTGCTGGCACTACCAACCAAGGAGTGTGGGTGTCTCC	1500
Db	1769	-----GGACCAAGAGGCCATATTTGCTGGCACTACGGCGCTCGAACCGTGGGGATGTGCTT	1823
QY	1501	GCGAAGACTGTGTGTGGCCCAAGTGTACTGTTTCACCCCAGCCCAAGTGTATGTGGGACG	1560
Db	1824	GCGTCGCAAGGTGTGTGGTCCAGTGTATTTGCTTCACTCCGAGCCCTGTGTGTGTGGGGACG	1883
QY	1561	ACCGACAGGCTTTGGAGGGCCCACTTACACGTGGGGGGAGATGTAGACAGATGTCTTCTTA	1620
Db	1884	ACCGATGCTTTGGGGCGCTCTTAAGTATAGCTGGGGGGAGATGTAGACAGATGTCTGTCTA	1943
QY	1621	TTGAACAGACTCGACCAACCGCTGGGGTCATGTTCGGCTGACAGGTGATGAACTCTTCT	1680
Db	1944	CTTAGCAACACGCGGGCGCTCTCAAGGCAACATGTTTGGGTGACAGTGGATGAACAGCACT	2003
QY	1681	GGCTAACACAGACTTGTGGGGCGCACACCTTCCGCTACTAAGCTGACTTCAAGCCACGC	1740
Db	2004	GGGTTTCACAAAGACGTGTGGGGGGCCCTCCGTGCAACATCGGGGGGGTCCGCAACAAACA-	2061
QY	1741	ACGGACCTGTGTGTCCCCCAACGGACTGTTTTAGAAGACATCTGATACACTTACCTTCAAA	1800
Db	2062	-----CTTTGTGTCTCCCAACGGAATTGCTCCGGAAGACACCCGAGGCCACTTACACAAG	2117
QY	1801	TGCGGCTCTGGGGCCCTGTGCTCAAGCCAAAGGTGCTGATGACTACATCCCTTACAGGCTTGG	1860
Db	2118	TGTGTGCTCGGGGGCCCTGTGTGACACCCAGGTGCATGTGTGACTACCATACAGGCTCTGG	2177
QY	1861	CATTACCCCTGSCACAGTTAATATATACCATCTTCAAAATAGAGATATGTGGAGGGGTT	1920
Db	2178	CACTAACCTGTGACTGTTTAACTTTAACGCTCTTTAAGTCAGAGATATGTGGGGGGGTG	2237
QY	1921	GAGCAACGGCTCAACGGCTGCATGTGCAATTTCACTCGTGGGGATCGTTGCAACTTGTGAGGAC	1980
Db	2238	GAGCAACAGGCTCAATGTGCTGATGCAATTTGACTCGAAGAGAGCGCTGTGACTTGGAGGAC	2297
QY	1981	AGAAACAGAAAGTCAACTGTCTCTCTTTGTGTGCACTCCACACAGGAATGGGCAATTTTACT	2040
Db	2298	AGGATATAGTTCAGAACTCAAGCCCGCTGTGTGTCTTCACAAAGATGGCAAGTACTGCC	2357
QY	2041	TGCTCTTAAGTGGAGCTGCGCGGCTGTGCACTGATCTTCTCCACTCCACCAAAACATC	2100
Db	2358	TGTTCTCTTCAACACCTTACCGGCGCTGTGCCACTGGCTTGATCAATTTCAACCGGAACATC	2417
QY	2101	GTGACGCTAACATTCATGATATGCGCTTATCACTGCGCTCACAAATACATGTGCGATGG	2160
Db	2418	GTGACGCTGCAATATCTGTATCGGTATAGGGTGGGACAGTGTCTCTTTGGCAATCAAAATGG	2477
QY	2161	GAGTGGGGAATATACCTTATTCCTGTCTTATACGGAGCGCAGGGTTTGGCGCTGTATGG	2220
Db	2478	GAGATATATCTGTGTCTTTCTTTCTTCTTGTGGGAGCGCGCGTGTGCTGTGGCTGTGG	2537
QY	2221	ATGTCTCATCTTTGTGGGCGCAGGCGGAAGACAGCACTAGAGAACTGTATCTTGTCAACGCT	2280
Db	2538	ATGATGTGCTGTGATAGCCCAAGGCTGAGGGCCACTTAGAGAACTGTGTGTCTCTCAATGGG	2597
QY	2281	GCGAGCGCACTACTACTGCATATGTGCTTCTATATTTTGTCACTTTTTCGTGGCTGTGG	2340
Db	2598	GCGTCTGTGTGCGGAGCGCATGTGGCTTCTCTCTCTCTCGTGTCTTCTGTGGCGCGCTGG	2657
QY	2341	TACATCAAGAGTCCGGGATAGTCCCTTATGCTATCTATTTCCCTCACTGGCTGTGTCTCTT	2400
Db	2658	TACATCAAAAGCAAGCTGTGTCTCTGTGGGCGGCATATGTCTCTATGTGGGTATGGCCGTG	2717
QY	2401	AGCTTACTGTCTCTTACAGATATGCCCAACAGGCTTATGTATGAGCGACTGTGTGATGGC	2460
Db	2718	CTCTGTGTCTTGTGGCCCTTACACACACAGACTTATGTCAATGACCGAGAGATGTGTGCA	2777

OY	2461	CAGATPAGAGGGGCTCTGCTGATATATCACTCTTTATCTCAACCCCGGATPAG	2520
Db	2778	TCGTCGGAGGGCGGGTTTTGTGAGTCTGTACTCTTGAACCTTGACCACTAATAAG	2837
OY	2521	ACCCCTTCAGACCGGTTTTTGTGGTGTGTGCTATCTTCGACCCCTGGGGAAAGTATG	2580
Db	2838	GTGTTCTCGCTAGGCTCATATGGTGTGTACAAATTTTATCACAAGCCGAGGCCAC	2897
OY	2581	GTCCAGAGTGGGCAACCACTATGACAGTGGCGGGTGGCCGTGATCATATGGGCC	2640
Db	2898	TTGCAAGGTGGTGGCCCCCTCAAGTTGGGGAGGCGCGATGCATCATCTTCCTT	2957
OY	2641	GTCCCAATTTCTAACCAAGGTGTGTGTTGACATACAAAGTGGCTCTGGCGGTCTT	2700
Db	2958	ACATGGCGGGTCCATCCAGAGCTAATCTTTGATCATCAACAAACCTCTGTGGCATCTC	3017
OY	2701	GGGCGTGGCTTACCTCTAAAGGTGCTTGTGACGGCGTGGCCGTACTTCGTGAGGGCTCAC	2760
Db	3018	GGTCGGCTCATGTGTCTCCAGGCTTGGGCAATACTAGATGCCGTACTTTGTACGGCTCAG	3077
OY	2761	GCTCTACTGAGGATGTGCACCAAGGCAAGGATCTGCGGGGGGCAAGTACGTACAGATG	2820
Db	3078	GGGCTCATCCGTGATGATGTATGTGGGGAAGGTGCTGGAGGCCATATGTCCAAATG	3137
OY	2821	GGCGTACTAGCCCTTGGGAGGTGCATCGGCACTTACATCTATGACCACTCACCCCTATG	2880
Db	3138	GCCTTCAATGAAGCTGGGCGCGCTACAGATACGTACATATATGACATCTTATCTCCACTG	3197
OY	2881	TGGATTTGGGCTGTAGTGGCTGCGGGGACCTGGCGGGTGGCCGTGAGCCATCATCTTC	2940
Db	3198	CGGATTTGGGCCCAACGGGGGCTTACGAGACCTTGGGTGGCATGTAGAGCCCGTGTCTTC	3257
OY	2941	AGTCCGATGAGAAAGAAATCATTTGTCTGGGGAGCCGAGACAGCTGTGTGTGGGACATT	3000
Db	3258	TCCTGACATGAGAGATTAATCTATATCACTGGGGGGGCAACAACGGGGGTGTGGGGAATCT	3317
OY	3001	TTTACACGAGCTTCCCGTGTGCGGCCGACATTGTGTGGGAGGTCTCTTGGCCCACTGAT	3060
Db	3318	ATCTCGGGTCTTACAGATCTCGCCCGAAGGGGGAAGATGACTTCTTATGAGACCGGCCGAT	3377
OY	3061	GGCTATPACCTCAAGGGGTGGAGTCTTGGCCCCCATCATCTGTTAAGCCGACGACACA	3120
Db	3378	AGTTTGGAGAGCAGGGGTGGCGGCTCTTGGCTATACAGGCTTATTTCCCAACAAACG	3437
OY	3121	CGTGGCCTTTTGGGCAACATATGTTGATGAGATGACGGGGGCGCGACAGACAGACAGGCT	3180
Db	3438	CGGGGCTCGCTGGCTGTATCATATCAGCTCACAGGTGGGACAAAGAACAGGTGAT	3497
OY	3181	GGGGAAATTCAGGTCTGTGTCCACAGTCACTAGTCTCTTCTCGAACATCATCTCGGGG	3240
Db	3498	GGGAGGTTGAGGTGCTCTCCACCGCAACCACTTTCTCTGGGGAACCTCGTCAATATGGC	3557
OY	3241	GTTTTGTGAGCTGTCTACATGAGAGCTGGGACAAAGACTTGGCGGGCTCACGGGGTCCG	3300
Db	3558	GTGTGTGGAGCCGTACACATGATGGTGGCGGCTCGAAGAACCTTGGCGGGCCGAAAGGATCA	3617
OY	3301	GTCAACGAGATGTACTCCAGTGTGAGGGGGGACTTATAGAGGTGAGCCCAAGCCCTCTGGG	3360
Db	3618	ATCACCCCAATGTACACCAATGTATGACACAGAACTGTGCGGCTGGCGGGCGCCCCGGG	3677
OY	3361	ACTAAATCTTTGAGACCGTGCACGTTGTAGAGCGGTGCAGCTGTATCCTGTGCACGGGAAAC	3420
Db	3678	GCGGCTTCATGACACGTTGCACCTGGGCACTGGAGCTTGAACCTTATCTTGATCAGAGGCAT	3737
OY	3421	GCTGATGTCAATCCGGGCTCGAAGACCGGGGACAAACGGGGAGCGCTATCTCTCCCGAGA	3480
Db	3738	GCTGATGTGTCTTCGGGTGGCGCCGGCGGGGGAGACAGCAGGGGGAGCCGTCTTTTCCCCAGG	3797
OY	3481	CCTTTTTCACACTTGAAGGGGTCTCTCAGAGAGCCCGGTGTATGTGCCCAAGGACAGCT	3540
Db	3798	CCCATCTCTTACCTGAAGGGGCTCTCGGGGTGACCACTGCTTTTGTCCCTTGGGGACAGTT	3857
OY	3541	GTCCGAGTCTTTCGGGGAGCTGTGTGTCTCTCGGGGCGGTGACTATGATCATATATTCATC	3600

Db 3858 GTAGGCACTCTCCGGGCTGTGTGTGACCCGGGGGTTGGAAAGGGGTGACATTGATA 3917
Qy 3601 CCCGTTGAGACATCGACATGTGACGCGGTCCCAACCTTTAATGACAACAGACACCA 3660
Db 3918 CCGGTGAGTGTATGAAACATCAATGGGTCTCCGGTCTTCAACAGCAACTATCCCT 3977
Qy 3661 CCGTGTGCCCCAGACCTATCAGTCGGGTACTTGTGATGCCCCGACTGGCAGTGAAG 3720
Db 3978 CCGGCGGTACCGCAAACTTCGAATGGACATTTAAGCGTCCCACTGGAGGGGAG 4037
Qy 3721 AGCACAAAGTTCTGTGTGCAATGTCTCAAGGGTATAAAGTGTAGTCTTAATCC 3780
Db 4038 AGACCAAAAGTGGCGGTGATATGACAGCCAAAGGTACAAAGTGTCTGTAAACCG 4097
Qy 3781 TCAGTGTGTCACCCCTGGGGTTTGGGGCTACTTGTCTAAGGACATGGCATATCC 3840
Db 4098 TCCGTTGCGGCACATTGGGCTTTGGAGCGTATATGTCCAAAGGACATGGCATGAGCT 4157
Qy 3841 AACATTAGACTGAGTGAAGATGTGTGACACCGGGGCGCCATCAGGTACTCCATAT 3900
Db 4158 AACATCAGAACTGGGGGTAAAGACCATCACACGGGGCGGCCCATCACTGACTCCACTAT 4217
Qy 3901 GGCAATTTCTCGCCGATGGGGGCTGTGCGGGGCGGCTTACGACATCATATGTGAT 3960
Db 4218 TGCAGTTCTTGGCGAGCGGTGATGTCCGGGGCGGCTTATGACATCATATATGTGAT 4277
Qy 3961 GAATGGCATGCGGTGACCTTACACCACTCCCTTGGCATCGGAAGCTCTTATCAAGCA 4020
Db 4278 GAAATGCCACTCACTGACTCGATCACTCTTGGGATCGGACAGATCTTGATCAAGCA 4337
Qy 4021 GAGACAGCTGGGGTTCAGATCTGTGTGCTCAAGCTACGCCCTTGGTCAGTACA 4080
Db 4338 GAGACGGCTGGAGCGCGGCTGTGCTGTGCTCCACCGCACGCTCCGGAGTCATACC 4397
Qy 4081 ACCCCCAACCCCAATAGAGAGGTGGCCCTTGGGACAGAGGCGAGATCCCTTCTAT 4140
Db 4398 GTGCGACACCCCAATGAGAGAGTGGCCCTGTTCACACTGAGAGATTCCTTCTAT 4457
Qy 4141 GGGAGGGGCGATTTCCCTGTCTTACATCAAGGGAGGAGACATGTCTTGTGCATTCA 4200
Db 4458 GGCAAAGCATTCCTCATTTAGAGCCATCAAGGGGGAGAGCATCTCATTTTGTGCATTC 4517
Qy 4201 AAGAAAAAGTGAAGACTCGCGCGGCGCTTGGGGTATGGCTTGAATCTAGTGA 4260
Db 4518 AAGAAAGATGTGAAGACTCGCGCAAGCTGACAGGCTCGGACTCAATGCTGTAGCG 4577
Qy 4261 TACTTCAAGAGGTGAGAGTCTCCGTAATACCACTCAGGGAGAGTGTGTGTGCGC 4320
Db 4578 TATTACCGGGGTCTGATGTGTCCGTATACCGACTAGCGAGAGTGTGTGTGCGCA 4637
Qy 4321 ACCGAGGCTCATAGCAGGGTATACTGGGACTTTGACTCGTGTGACTGACTGAAAGCTA 4380
Db 4638 ACAGAGCTCTATATACGGGTTTACCGGCGACTTTGACTCAGTGTATGACTGAAACCA 4697
Qy 4381 GCGGTACTCAAGTTTGAAGCTTCAAGTTAAGCCCACTTACCATAACCAAGATT 4440
Db 4698 TGTGTACCCAGACAGTGGATTTCAGTTGATCCCACTTACCATTAAGAACCAAG 4757
Qy 4441 GTCCCTCAAGACGTGTCTCAAGTAGCAGAGCGGGGTCCGACGGGTAGGGAGAACTG 4500
Db 4758 CTGCCCCAAGACGCGGTGTGCGGTGCGAGCGGCAAGGTAGGACTGAGGGAGGAGAT 4817
Qy 4501 GGCAATTAAGTATGTTTCCACTGTGAGCGAGCTCAGGAATGTTTGAAGTGTAGT 4560
Db 4818 GGCACTTACAGGTTTGTGATCTCAGAGAAACGGCCCTCAGGCAATTTGATCTCTCGGTC 4877
Qy 4561 CTGTGTAGTGTCTAGACGAGGGGCGCATGTATGAGCTCACCACTCGAGAACCA 4620
Db 4878 CTGTGTAGTGTCTAGACGAGGCTGTGTGTATGAGCTACGCGCGGTGAGACCTCG 4937
Qy 4621 GTGAGGCTCAGGGGTATTTTCAACACGCGCGTTTGTGTGTGCAAGACCATCTTGAG 4680

Db 4938 GTTAGGTTGCGGGCTTACCTAAATACACAGGGTTGCCGCTGTGCAGAACCACTTAG 4997
Qy 4681 TTTTGGAGAGCATTTTCAACCGGCTCACACATAGATAGCCACTTCTTCCAAACA 4740
Db 4998 TTTTGGAGAGGCTTCTTCAAGGCTTCAACCCATATAGATGCCACTTCTTGTCCAGACC 5057
Qy 4741 AAGCAATCGGGGAAAATTTGCGATATCTTAACAGCTTACAGGCTTCAAGTGTGCGCTAG 4800
Db 5058 AAACAGGAGAGAACACTCCCTTACTGTATAGCATTAACAGCCACAGTGTGCGCAG 5117
Qy 4801 GCCAAGCCCCCCCCCGTCTGGGACGTATGTGTGAAGTGTGTTGCTGACTCAAGGCC 4860
Db 5118 GCTCAGGCTTCACCTTCATGTGTGACCAAAATGTGAAGTGTCTATAGCTTAAAGCCC 5177
Qy 4861 ACACGTGTGGGCCCAACCTCTCTGTATCCGCTTGGGCTGTGTTACCAACAGAGTCA 4920
Db 5178 ACACGTGATGGGCAACGCGCTGTGTACAGGCTTAGAGCGGTTCAAAATAGGTCAT 5237
Qy 4921 CTCACATCCGTGACGAATATCATGCGCACTGTATGAAACCGCACTTGAAGTCATG 4980
Db 5238 CTCACACCCCATTAACAAATATCATGAGCATGATGTGCGCTGACCTGAGAGTGTCTC 5297
Qy 4981 ACCAGCAATGGGTCTTGGCAGGGGAGTCTTGGGCGCGCTGCGCGGATTCCTGAGG 5040
Db 5298 ACTAGCACTTGGGTGCTATGAGCGAGTCTTGGGCTTGGGCGGCTGACTGTCTGAGC 5357
Qy 5041 ACCGGGTGTGTGATCATTCGCGCGCTTGCACATTAACAGACGCGCTGTGGCGCG 5100
Db 5358 ACAGGAGGCTGTGATTTGTGGGCAAGATCATTTGTCCGGAGGCGACGTGTATTC 5417
Qy 5101 GACAGAGGCTCTTATAGAGCTTTTATGATGAGAGAAATGTCTCTTAGGCGGCT 5160
Db 5418 GACAGGAAAGTCTCTTACAGAGAGTTCATGATGAGTGAAGAGTGTCTTCAACCTCC 5477
Qy 5161 CTGATTGAAGAGGGGAGAGGATAGCGGAGATGCTGAATCCAAATCCAAAGCTTATG 5220
Db 5478 TACATGAGCAGAGAAATGCAAGCTCGCGAGCAATTAACAAGAGCGCTCGAATTCG 5537
Qy 5221 CAGCAAGTTCCAACAGCTCAAGACATACCAACCACTGTGAGGCTTATAGGCCCAAG 5280
Db 5538 CAACAGCAGCAACAGCAAGAGGAGCTGTGCTCCGCTGTGGAGTCCAAAGTGGCAGCC 5597
Qy 5281 GTTGAACAATTTGGGCAACACATGTGAACTTATTAAGGACATCAATACCTTCGA 5340
Db 5598 CTGAGAGTCTTGTGGCGGAACACATGTGAATCTTCATGAGGAGTACAGTACTTGGA 5657
Qy 5341 GGACTATCAACCTGTCAAGGAAACCTGAGATGCTTCATGATGGCTTCAAGTGGCGCC 5400
Db 5658 GGCCTATCACTGTGCTGGAACCCCGATAGCATTTGATGCTTTTACAGCTCTT 5717
Qy 5401 CTCACAGTCCGCTGTCAACAGACCACTATCTTCTCAACATTTTGGGGGCTGTGCTA 5460
Db 5718 ATCACCAGCGCTGTACCAACCCAAATACCTCTGTTTAATCTTGGGGGATATGGTG 5777
Qy 5461 GCATCCCAATTGACACACCCGCGGAGCACTGCTTGTGTGTCAAGTGGCTTAGTGGGA 5520
Db 5778 GCTGCCCACTCGCTCCCGCAGCGCTGTGCTTGTGTGGCCCGCAATTTGCGGT 5837
Qy 5521 GTTCCGTAGGAGTATAGGCTTAGGTAAGTGTCTAGTGAACATCTGTGAGGGTATGG 5580
Db 5838 GCGGCGTTTGGACGATATGCTGTGGAAGGTACTTGTGACATTTTGGCGGGCTATGG 5897
Qy 5581 GCGGCAATTTGCGGGGCTCTCGTCAATTCAGATCATGTCTGGCAGAAAGCCCTTCATG 5640
Db 5898 GCGGGGGTGTGTGGCCACTCGTGGCTTTAAGTATAGAGGGGAGATGCGCTCACT 5957
Qy 5641 GAGAGTGTCTCAACTTGTGCTTGAATTTGTCTTCCGGGTCTTGTGTAGTGGAGTC 5700
Db 5958 GAGGATCTGTTAATTTACTCTCTGCAATCTTTTCTGTGGCGCCCTGTGTGTGGGGTCTC 6017
Qy 5701 ATCTGCGCGGCAATTTGTGCGGCAACGTGTGGAGCGGGGGAAGGCGCGTCCAAATGGA 5760
Db 6018 GTGTGCGACAAATACGTGTGTGCGACGTGGGCCCGGAGAGGGGCTGTGAGTGTGAG 6077

Qy	5761	AATGACATCAATGGCTCTTGGCTTCCAGAGGAAATACAGTCGCCCCCAACCACTACTGACG	5820
Db	6078	AACCGCTGAATAGGCTTCGCTTCGGGGGATACACAGCTCCCCACGCACTAATGTGCG	6137
Qy	5821	GAGTCGATGCGTGCAGCGCTGTGACCCAACTACTGGCTCCCTTACCATTAACAGCGCTG	5880
Db	6138	GAGAGCGACGCGCGGGCGCGTGTTAATCTCAGATCCCTCTCCACCTTACATCACTCAAGTTG	6197
Qy	5881	CTCAGAAAGCTCCCAACTGGATTATCTAGAGACTGCCCCATCCATGGCGCGCTCGTGG	5940
Db	6198	CTGAAGAGGCTTCATCAGTGGATTATATAGAGACTGCTCCACGCTTGTTCGGGCTCGGG	6257
Qy	5941	CTCCGCGATGTGTGGAGCTGGGTTTGGACCAATCCCAACAGACTTTAAATAATTTGGCTGAC	6000
Db	6258	CTAAGAGATGTTTGGGACTGGAATATGACGSGTGTGAGTACTTCAGAACCTTGGCTCCAG	6317
Qy	6001	TCCAAATTAATTCCTCAAGAAAGTCCCGCGCTCCCTTGTCTCTGTCAAAAAGGGGTACAAG	6060
Db	6318	TCCAAGCTCTCGCGCGGTTACCGGAGACTCCCTTCTCGTATCTCAACGCGGGTACAAG	6377
Qy	6061	GGCGTTGGCGCGGCGCACTGGCATATGACCAAGGTTGCTTTCGCGCGCAATATCTCT	6120
Db	6378	GGAGTCTGGCGGGGAGATGGCATATGAAACCACTGGCCATGTGGAGCAATATCAAC	6437
Qy	6121	GGCAATGTCGCTTGGGCTCCATGAGAATCACGGGGCGTAAAGCCTGATGAATATCTGG	6180
Db	6438	GGACATGTCAAAAATGGCTCCATAGAGATTGTGGCGCAAAAACCTCGACGAAACGTTGG	6497
Qy	6181	CAGGGAGCTTTCCTATCAATTTGTTACAGGAGGGCGAGTGCGTGGCCGAAACCCGCGCA	6240
Db	6498	CATGGAACATTTCCCATCAACGCAATACCAACGCGGCCCTCGCACGCCCTCCCGACGCGG	6557
Qy	6241	AACTTAAGTGGCATCTGAGGGGTGGCGGCTCAAGATACGCGGAGGTGACGACGAC	6300
Db	6558	AACATATTCAGAGGCGCTGTGGCGGGTGTCTGTGAAGATACGTGAGAGTTTACCGGGGTG	6617
Qy	6301	GGGTCAATACCACTATTAACAGGACTCACTCACTGATTAATTGAAGTCCCTGCAACTA	6360
Db	6618	GGGATTTCCACTACGTACGCGGATGACCACTGACCAACGTGAATGCCATGCGCAAGTT	6677
Qy	6361	CCCTCTCCCGAGTTCTTTTCTGGGTGAGCGGATGCAATTCATAGTTTGGCCCCACA	6420
Db	6678	CCAGCCCCCTGAATTTTTCACGAGAGTGAATGAGTACAGTTGACAGGTAATGCTTCCAGTG	6737
Qy	6421	CCGAGACGTTTTTCCGGAGTAGAGTCTCGTTCTGCGTGGGCTTAAATTCATTTGTGCGT	6480
Db	6738	TGCAAACTCTCTCTACGAGAGAGAGTGTATTCCAGTTCGGGCTCAACCAAGTACTGGTCT	6797
Qy	6481	GGGTCCAGCTTCTCTTGGCACTCTGAACCCGACACAGACGTAATGATGTCCATGCTAACA	6540
Db	6798	GGGTCAACAGCTCCATGTGAGCCGGAACCGGATGTGGCAGTGCTCATCTCATCTCAC	6857
Qy	6541	GATCATCTCTATCACGCGGGAATCTGACGCGGGGCGTTTAAAGCGGGGCTCACCCCCA	6600
Db	6858	GACCCCTCTCATATTACAGACAGAACCGCCAGGTAAGCTGGCTGCACGAGGAGGTCTTCCCCC	6917
Qy	6601	TCCGAGGCAAGCTCTCGGCGAGCGAGCTATCGGACCACTGCGTGCAGGCACTGCAAC	6660
Db	6918	TCTTTGGCAGCTCTTCAAGCTTAAGCAGTGTGTCTGCGGCTTCTTGAAGGCGACATGTAACT	6977
Qy	6661	ACCCAAGCAAGCTTATGATGTGACATGTGTGAATGCTAACT-----GTTT	6708
Db	6978	ACCCATCATGACTCCCGGAGGCTGAACCTCATCGAGGCAACCTCTGTGGCGGACAGAG	7037
Qy	6709	ATGGGAGGCGATGTGACTCGGATAGAGTCTGGGTTCAAAAGTGTGTTCTTGACTTCTC	6768
Db	7038	ATGGGCGGGAATCAACCCGTGTGAGATCAGAAATAAGGTGTATCTGTGACTTTTTC	7097
Qy	6769	GACCAATGTGTGAAGAAAGAGCGACTTGAAGCTTTCGATACATCAAGATACATGCTC	6828
Db	7098	GATCCGATTGGGCGGTGAGAGATGAGAGGAAATATCCGTCGCGGAGATCTGTGGA	7157

6829 CCGAAGAGGGTTCCCAACGCTTTAACGGCTGGGACGGCGTGAATTACAAACCAAG 6868
Db AAACCAAGAAAGTTCCCCACAGCGTTGCCATATGGGACAGCCGGAATTACAACTTCCA 7158
Qy CTTTGGAATCGTGGAAAAAGCCAGATTACCAACGGGCACTGTTGCGGGCTGTGCTTC 6889
Db CTGCTAGAGTCCGTGGAAAGAACCCGGACTACGTCGCCCGGTGGTACACGGGTCCCTTGG 7218
Qy CCTCTCTTAGGAAAAACCCCGACGCTTCCCAAGAGGCGCCGACAGTGGGCTTAAGT 6949
Db CCATTTACCAAGGCCCCCAATACCACTCCACGGAGGAAAGAGCGGTGTCTGACA 7278
Qy GAGGACTCATAGAGATGGCCCTTGAACAGCTGGGCATTAAGTCCTTTGGCCAGCCCCC 7009
Db GAGTCACCGGTCTTCTTGCTTGGCGAGCTCGCTACTTAAGACTTTTGGCAGCTCCGG 7338
Qy CCAAGCGGCGAATTACAGCCTTTCCACGGGGGCGGGCGTGCAGATTCCGGCAGTACAG 7069
Db TCGTGGCCGTGTAACAGCG---CAGCGGAGCTGGCCCTCCGA-----TCAGGCC 7398
Qy CCTCTGATGAGTTGACCTTTGGAGACAGGTTCCATCTTTCATGCCCCCTCGAG 7129
Db TCCGACGACGGGCGAACAAAGATCCGACGTTGATGTACTCTCCCATGCCCCCTCGAG 7446
Qy GGGGAGCTTGGAGATCCAGACTCGAGCCTGACAGGTAGACAGTACGCCCAACCCCCCAG 7189
Db GGAGAGCCAGGGAGACCCGACCT----- 7506
Qy GGGGGGGTGGCAGCTCCGGGCTCGGACTCGGGGCTCTGGTCTACTTGGCTCCGAGGAGAC 7249
Db -----CAGCGACGGGTCTTGGTCTAACGTTGACGGGAGGAGCTGGT 7529
Qy GACTCCGTGTTGTGCTGCTCATATGTCATATCTCGGACCGGGGCTCTAAATCTCTGT 7309
Db GAGGACGTGTCTGCTGCTCATATGTCATATATGGAACAGGTGCTTGAATCAGCATGC 7569
Qy AGTCCCGAAGAGAGAGATTACCGATTAAACCCCTTGGACACTCCTGTTGGATATCAC 7369
Db GCTGGGAGGAGAGCAAGTGGCCATCATACCTGTTAGCAACTTTGTGCTGCATCACAC 7629
Qy AACAGGTGTACTGTACCAACAACAAAGCGGCTCATTAAGGCTTAAAAAGGTAATTTT 7429
Db AGTATGTCTACTCCACAACATCTCGACGGGCAAGTCTGGGCGAAGAAAGGTCACCTTT 7689
Qy GATAGGATGCAAGTCTGACTCTACTACTACGACTCAGTCTTAAAGGACATTAAAGTAGCG 7489
Db GACGACCTGCAAGTCTTGGAGCAACACTACGGGACGTCTCAAGAGATGAAAGCGGAAG 7749
Qy GCCTTCAAGGTACCGCAAGGCTCTCAACATGAGAGAGGCTTGCAGATTAAACCCGACC 7549
Db GCGTCACAGTTAAAGGCTTAGGCTTCTATATAGAGAGAGGCTGCAAACTGAGCGCCCCA 7809
Qy CATTTGCAAGATCTAAATATGGGTTTGGGGCTTAAGAGAGTCCGACGCTTGTCCGGAAG 7609
Db CATTCGGCAATCCAAATTTGGCTACGGGGGGAAGAGTCCGGGCTTATCCACAGG 7869
Qy GCCGTTAACCAATCAAGTCGTTGGAGAGGACCTCTTGGAGGACCTCAGAAACCAAAATT 7669
Db GCCGTCAACACATCCGCTCGGTGGGAGGACTTGTGGAAAGACCTGAAACCAAAATT 7929
Qy CCGACAAACATTATGCGCAAAATAAGAGTGTCTGCGTGGACCCCAACAGGGGGCAAG 7729
Db GATACCAACATCATGCGCAAAATAAGGTTTTCTGGGTCCAAACAGAAAGAGGCGCG 7989
Qy AAAGCAGCTCCGCTTATGTTTACCTTGAACCTTGCCTCGCGCTCAGGGTCTGCGAGAAAGATGCC 7789
Db AAGCAGCTCCGCTTATGTTTATCCAGACCTGGGGGTACGTGTATGCGAAGAAAGATGCC 8049
Qy CTTTATGACATTACCAAAAACTTCTCTCAGCGGGTATGGGGGCTCTTATGATTTCCAG 7849
Db CTTTACGACGTGGTCTCCACCTTCTCAGGCGCGTGTATGGGCGCCCTCTCATACGATTTCCAG 8109
Qy TATTCGCCCGCTCAGCGGGTAGAGTTTCTCTTGAAGACATGGGCGGAAAAAGAGACCTT 7909

Db	8169	TACTCTCTGGGCAAGCGGGTGAAGTTCTCTGGTGAATACTGGAAATCAAAGAAATGCCCT	8228
QY	7969	ATGGGTTTTTGGATGATATCCCGATGCTTTTGACTCAACCGTACTAGAGAGACATCAG	8028
Db	8229	ATGGGCTTCTCATATAGACACCCGCTGCTTTGACTCAACGGTCACTGAAATGACATCCGT	8288
QY	8029	ACTAGGAGTCCATATATCCGGCCCTGCTCCCTTGGCCCGAGAGAGCCCACTGGCATACAC	8088
Db	8289	ACTGGGAATCAATTTACCAATGTGTGACTTGGCCCCGAAAGCCAGGAGCCCATTAAG	8348
QY	8089	TGCTAACTGAGAGACTTTAGCTGGGAGAGGCGCTATGTTCAACAGCAAGGCGCAAACTGC	8148
Db	8349	TGCTCAACAGAGCGGCTTTATGTCCGGGGGTCCCCGTAATATGGAAGGGCAGAACTGC	8408
QY	8149	GAGTACAGAGCGTTGCGCGCCAGCGGGGTGCTCACCACTAGCATGGGGAACCATCA	8208
Db	8409	GGTTATCGCCGGGTGCGCGCAAGTGCGGTCTGACGACTAGCTGGCGCAACCTTCACA	8468
QY	8209	TGCTACGTGAAGCCTTAGCGGCTTTGTTAAGTGTGAGGGATATGCGCCCAATGCTG	8268
Db	8469	TGTTACTTGAAGGCCACTGCGGGCCGTGAGCTGGAAAGCTTCAGAGCTGACGATGCTC	8528
QY	8269	GTATCGGCGGATGACTTGTTGTGCTCATCTCAGAAAGCCAGGGGACCGAGAGAGACGAGCG	8328
Db	8529	GTGAACGAGAGAGACTTGTGTGTTATCTGTGAGATGCGGGAAACCCAGAGAGATGCGGCG	8588
QY	8329	AACTGAGAGCCTTTCACGAGAGGCTATGACAGATATTTGCCCCCTCTGTGTGACCCCCC	8388
Db	8589	GCCCTACGAGCCTTACCGAGGCTATGACTAGTATTCGGCCCCCGGGGAGACCGGCC	8648
QY	8389	AGACCGGAGTATGATCTGTGAGCTGTAACATTTGCTCTCAATATGTCTGTGGCGCTG	8448
Db	8649	CAACGAGAATACGACTTGGAGCTGTATACGTATCTCCTCCAAATGTGTGGTCCGCGAC	8708
QY	8449	GGCCCAAGAGCGCGCGCGAGACTACTACCTGACACAGAACCTTACACTCCAATCGCGCG	8508
Db	8709	GATGATCCGGCAAAAAGGTGTACTACTCTACCGGTGACCCGACCAACCCCTTGCACGG	8768
QY	8509	GCTGCTGGGAAACAGTTTAGACACTCCCTGTCAATTTGATGCTGGGAAACATCATCCAG	8568
Db	8769	GCTGGTGGGAGACAGTTTAGACACTCCAGTCACTCTGTGCTTAGGCAATATCATATG	8828
QY	8569	TACGCCCCGACCATATGGGCTTCGATGCTCTGATGACACACTTCTTCCATTTCTCATG	8628
Db	8829	TATGCGCCCACTTATGGCGGAGATGATTCGATGACACTCAATTTCTTCTATCTTCTTA	8888
QY	8629	GCTCAAGACGCTGAGCCAGAACCTTCAACTTTGAGATGTACGAGCGGTGATCTCCGTG	8688
Db	8889	GCTCAGGAGCAACTTGAAAAGCCCTGATGTGCAGATCTAAGGGGCTGTATCTCCAT	8948
QY	8689	AGTCCCTTGGACCTCCCAAGCTATATTTGAAAGTTACATGGGCTTGAAGCTTTTCTCTG	8748
Db	8949	GAGCCACTTGAAGCTTACCTCAGATCATTTGAAAGCATCTCAATGCTTTAGGCGATTTTCACTC	9008
QY	8749	CACACATACCTCCCAACGACTGACACGGGTGGCTTACGCCCTCAGAAAATTGGGCG	8808
Db	9009	CACAGTTACTCTCAGAGTGAGATCAATAGGGGTGGCTTATGCTCAGAGAAATTGGGGTA	9068
QY	8809	CCACCCCTCAGAGCGTGGAAAGCCGGGAGACGTGACATCAGGGCGTCCCTCATCTCCCGT	8868
Db	9069	CCGCTTTTGCAGGTCTGGAGACATGGGGCAAAAGTGTCCGGCTAAGCTATGTGCCAG	9128
QY	8869	GGGGGGAGAGCGGCGTTTGGCGGTGATATCTCTCAATTTGGGCGGTGAAAGCAAGCTC	8928
Db	9129	GGGGGGAGGGCTGCGCACTTGGCGCAAGTACTCTTCACTTGGGCGAGTAAAGCAAGACTT	9188
QY	8929	AAACTCACTCAATGCGGAGACCGGCGCTCTGGAATTAATCCAGCTGTGTTACCGTCCGC	8988
Db	9189	AAACTCACTCCAAATCCGGCTGCGTCCAGCTAAGACTTGTCCGGCTGGGTTGTTGCTGGT	9248
QY	8989	GCGGCGGGGGGACATTTATACAGCGTGTGCGTGCCTGACCCCGCTTATTTGCTCTTT	9048

Db 9249 TACAAAGGGGAGACATATATCAAGCGCTGTCTGTCGCCGACC CGTGTGCATCATGTTG 9308

Oy 9049 GGCTACTCCTACTTTTGTAGGGGTAGGCTTTTCTACTCCCGGCTGG 9099
|||||
Db 9309 TGCTACTCCTACTTTCTGTAGGGGTAGGATCTACTGTCTCCCAACCGG 9359
|||||

RESULT 13
AAD49655
AAD49655 standard; DNA, 9413 BP.

XX AAD49655;
DT 24-MAR-2003 (first entry)
DE Hepatitis C virus (HCV) DNA.
KM Amyloidosis; haemophilia; Alzheimer's disease; atherosclerosis; cancer;
KM gigantism; dwarfism; hypothyroidism; hyperthyroidism; cystic fibrosis;
KM autoimmune disorder; aging; inflammation; diabetes; obesity; anorectic;
KM neurodegenerative disorder; Parkinson's disease; gene therapy; virocidase;
KM haemostatic; antibacterial; nootropic; neuroprotective; cytostatic;
KM fungicide; Hepatitis C virus; HCV; gene; ds.
XX OS Hepatitis C virus.
XX PN WO200283953-A1.
XX PD 24-OCT-2002.
XX PF 11-APR-2002; 2002WO-US011757.
XX PR 11-APR-2001; 2001US-0282965P.
XX PA (PTCT-) PTC THERAPEUTICS INC.
XX PI Rando R, Welch E;
XX DR WPI; 2003-075561/07.
XX PT Identifying a test compound that binds to a target RNA molecule for
XX treating or preventing amyloidosis, hemophilia, cancer, gigantism,
XX diabetes, by contacting a detectably labeled target RNA molecule with a
XX library of test compounds.
XX PS Example; Page 64-68; 152pp; English.
XX CC The invention relates to a method for identifying a test compound that
XX binds to a target RNA molecule, which comprises contacting a detectably
XX labelled target RNA molecule with a library of test compounds under
XX conditions that permit direct binding of the labelled target RNA to a
XX member of the library of test compounds so that a detectably labeled
XX target RNA:test compound complex is formed. The method is useful for
XX screening libraries of compounds for those that are selectively bind to a
XX pre-selected target RNA. The compounds are useful for inhibiting the
XX formation of a specific bound RNA:host cell factor complexes in vivo.
XX They are also useful for treating or preventing diseases associated with
XX overproduction or decreased protein function, such as amyloidosis,
XX haemophilia, Alzheimer's disease, atherosclerosis, cancer, gigantism,
XX dwarfism, hypothyroidism, hyperthyroidism, autoimmune disorders, aging,
XX inflammation, cystic fibrosis, diabetes, obesity, neurodegenerative
XX disorders, Parkinson's disease or infections (bacterial, viral, fungal).
XX The invention is also used in gene therapy. The present sequence is
XX Hepatitis C virus (HCV) DNA. This sequence is used to illustrate the
XX method of the invention
XX SQ Sequence 9413 BP; 1886 A; 2816 C; 2694 G; 2017 T; 0 U; 0 Other;

Query Match 45.5%; Score 4141.8; DB 8; Length 9413;
Best Local Similarity 66.6%; Pred. No. 0;
Matches 6066; Conservative 0; Mismatches 2952; Indels 93; Gaps 6;

1 ATGAGCACAAATCTTAACCTCAAGAAAACAAAAAACAACCAACCGTGCCCAACAA 60

Db 330 ATGAGCAAAATCTTAAACCTCAAAAGAAAAACAAGTAACACCAACCCGCGCCACAG 389
Qy 61 GAGCTTAAGTTTCGGGCGCGCGCCAGATGTTGGCGAGATATCTTTGGTCGCGCGAG 120
Db 390 GAGCTTAAGTTTCGGGCGCGCGCGCCAGATGTTGGCGAGATATCTTTGGTCGCGCGAG 449
Qy 121 GGGCCAGGTTGGGTGGCGCGCAAGAGAACTTCGAGCGGTCGCGCGAGCGTGG 180
Db 450 GGGCCAGGTTGGGTGGCGCGCAAGAGAACTTCGAGCGGTCGCGCGAGCGTGG 509
Qy 181 AGGCGCAGCCCAATCCCTAAAGATCGCGCTCACTGGCAAAATCTGGGAAAAACAAGA 240
Db 510 AGGCGCAGCCCAATCCCTAAAGATCGCGCTCACTGGCAAAATCTGGGAAAAACAAGA 569
Qy 241 TACCCCTGGCCCTTATACGGGAATGAGGAGACTCGGCTGGGCGAGATGGCTCTCCCC 300
Db 570 TACCCCTGGCCCTTATACGGGAATGAGGAGACTCGGCTGGGCGAGATGGCTCTCCCC 629
Qy 301 CGAGGTTCCGCTCTTGGGGGCGCCAAATGACCCCGGCAATAGGTGGCGCAAGTGG 360
Db 630 CGTGGCTCTGGCTTAATGGGAGCCCAAGAGCCCGGCGTGGGTGGCTTAATGGGT 689
Qy 361 AAGGTCAATGATACCTTAACGTGGGCTTTGCCGACTATGGGCTACATCCCTGTGCTG 420
Db 690 AAGGTCAATGATACCTTAACGTGGGCTTTGCCGACTATGGGCTACATCCCTGTGCTG 749
Qy 421 GGGGCCCCGCTCGGCGCGCTGCGCAAGCTCTCGCGCATGGCGTGAAGTCTGGAGAG 480
Db 750 GGGGCCCCGCTCGGCGCGCTGCGCAAGCTCTCGCGCATGGCGTGAAGTCTGGAGAG 809
Qy 481 GGGGTTAATTTTGAACAGGGAATTTACCCGGTTGCTCTTTTCTATCTTCTGGTGGCC 540
Db 810 GGGGTTAATTTTGAACAGGGAATTTACCCGGTTGCTCTTTTCTATCTTCTGGTGGCC 869
Qy 541 CTGCTGTCTGATCAACACCCCGGTCTCGCTGCGAGATGAAGAACATCAGTACCGGC 600
Db 870 TTGCTGTCTGTTGAACATCCGAGCTTCGAGTTAGAGAGTGGCAACGTGTCCGGGATA 929
Qy 601 TACATGTGATCAACGACTGACCAATGACAGATTAATCTGGAGCTTCAGGCTGTCTG 660
Db 930 TACATGTGATCAACGACTGCTCTCAACTCAAGTATGTGATGAGGCGAGCATGATC 989
Qy 661 CTCACGTCCTCCGGGTCGCTCCGTTGGAGAAAGTGGGAAATGATCTCACTGTGGATA 720
Db 990 ATGCAACCCCGGAGTGGCTGCTGCTCGGAGAGTAATTTCTCCGTTGCGGATA 1049
Qy 721 CCGGTCACACGATGTGGCGCTGCAAGCGGCCGCGCTCAACGAGGCTTGGCGAGC 780
Db 1050 GCGCTACCTCCACGCTCGCGGCAAGAGACAGAGCATCCCAACACAGAAATACAGAGC 1109
Qy 781 CACATGCAATGTTGATGTGCGGCAAGCTGTGCTGTGCTCTTACGTGGGAGACTTC 840
Db 1110 CAGGTGATTTGCTGCTGGGCGGCTGCTCTGTTCCGTAATGATGTTGGGAGATCTC 1169
Qy 841 TGGGTTGGGAGTGTGCTCGAGCCCAAAATGTTTCAATTTGCTGCGCGAGCACTGCTTT 900
Db 1170 TGGGTTGGGAGTGTGCTGCTGCTGCTGCTGCTTCACTTCTCACTGCGCCGATAGAGC 1229
Qy 901 GTCCAAAGATGCAATGCTCAATCTACCTTGTACATCACTGACACGCGATGGACATG 960
Db 1230 GTCCAAAGATGCAATGCTCAATCTACCTTGTACATCACTGACACGCGATGGATGG 1289
Qy 961 GAAATATGATGAATGAGTGGCCCAAGCTACATGATCTTTGGGCTTACGAGATGCGTGT 1020
Db 1290 GAAATATGATGAATGAGTGGCCCAAGCTTACATGATCTTGGGCTTACATCTCCGATC 1349
Qy 1021 CCGGAGTCAATTAAGACATCAATTAAGGAGGCTCATTTGGGAGCTCATTTGGCTGGCC 1080
Db 1350 CCAAGAGCGTGTGAGCATGTGGCGGAGCCCACTGGGAGTGTCTTAAGGAGGCTTGGCC 1409
Qy 1081 TACTTCTTATGACGAGGAGCGTGGAGAAAGTGTGTCAATCTTCTGTTGGCCGCGG 1140

Db 1410 TACTATTCATGTTGGGAACTGGGCTAAGGCTTATGATGATGCTACTTCTTGTGCG 1469
Qy 1141 GTGAGCGCGGCAACCTTACTGTTGGGGGTTTTCGCCGAGACCAACCGGCGCTCAAC 1200
Db 1470 GTTGAACGGGCAACCCACGTGACAGAGGAGGATGACTTCACAGACCCAGAGCTGTG 1529
Qy 1201 AGCTTATTTGACATGAGGAGCCAGGAGAAATCAAGCTGTCTTAACCAATAGCAGCTG 1260
Db 1530 TCTGGCTCTTCACAGGCGCATCTCAGAAATCACTGTAACCAACCAAGAGCTGTG 1589
Qy 1261 CACATCAACCGACCGGCTTGAACCTGATGATCTCTTGTGACACCGGCTTTATCGCTT 1320
Db 1590 CACATCAACCGACCGGCTTGAACCTGATGATCTCTTGTGACACCGGCTTTATCGCTT 1649
Qy 1321 CTGTTCTACACCCAGCTTCACTGTCAGATGTTCGGAACGCAATGTCCGCTGCGC 1380
Db 1650 CTGTTCTACACCCAGCTTCACTGTCAGATGTTCGGAACGCAATGTCCGCTGCGC 1709
Qy 1381 AGTATGAGGCTTCCGGGTGGGATGGGAGCGCTTGGCAATATGAGATTAATGTACCAAT 1440
Db 1710 CCGATGATGATGTTGCTCAGGAGGTGGGATCCATCACTATGATATGCTTGAAGCTC- 1768
Qy 1441 CCAAGGATATGAGACCTTATTTGCTGCACTTACCAACCAAGGAGTGGCGTGTCC 1500
Db 1769 -----GAGCCAGAGGCCATATTTGCTGGCATACGCGCTCGACGTTGCGGATGTGCTC 1823
Qy 1501 GCGAAGCTGTGTGGCCGAGTGTACTGTTTCAACCCAGCCAGAGTGTGAGTGGGAGC 1560
Db 1824 GCGTGGAGGTGTGTGCTCAGTGTATTTGCTTCACTCGAGCCCTGTTGTATGAGGAGC 1883
Qy 1561 ACCGACAGGCTTGGAGGCGCCACTTACACGTGGGAGGAGATGACAGATGTCTTCTTA 1620
Db 1884 ACCGATGTTTGGCGCTCTCTAGTATAGTGGGAGGAGATGACAGAGCTGTCTTA 1943
Qy 1621 TTGAACAGACTGACCAACCGTGGGCTCATGTTGGCTGACAGTGTAACTTCTTCT 1680
Db 1944 CTTAGCAACAGCGGCGCTCAAGGCACTGGTTTGGGTGACAGTGTAAACAGACT 2003
Qy 1681 GGTACACCAAGACTTGGCGGCGACCACTGCGCTGACTGACTGACTTCAACGCGAGC 1740
Db 2004 GGGTTCAACCAAGAGTGGGAGGCGCTGCTGTGCAACATGCGGAGGCTGGCAACAA- 2061
Qy 1741 ACCGACCTGTTGTGCCCAAGACTGTTTGAAGAGCATCTGATACCTTACCTTCA 1800
Db 2062 -----CTTGTGTGCTCCCAAGATGCTTCCGAGAGACCCAGAGGCTTACACAAAG 2117
Qy 1801 TGGGCTCTGGGCGCTTGGCTCACGCGCAAGGTGCTGATGACTTACCCCTTACAGGCTTGG 1860
Db 2118 TGTGGCTCGGAGGCGCTTGTGACACCAAGGTGATGATGATACCATACAGGCTTGG 2177
Qy 1861 CATTACCCCTGCACTGTAACTTATCACTTCAAAATGAGATGATGAGGAGGCTT 1920
Db 2178 CACTACCCCTGCACTGTAACTTATCACTTCAAAATGAGATGATGAGGAGGCTT 2237
Qy 1921 GAGCAGAGGCTCAAGGCTGATGCAATTTCACTGTGGGAGTGTGCACTTGAAGAGC 1980
Db 2238 GAGCAGAGGCTCAAGGCTGATGCAATTTCACTGTGGGAGTGTGCACTTGAAGAGC 2297
Qy 1981 AGAGACAGAGTCACTGTCTCTTGTGTGACTTCAACAGAGATGGGCAATTTTACT 2040
Db 2298 AGGAGATGATCAAGACCCGCTGCTGTCTTCAACAGAGTGGGAGATGCTGCC 2357
Qy 2041 TGTCTTATCTGAGACCTGCGCGCTTGTGAGCTGTGCTTCAACCTCCACCAAACTATC 2100
Db 2358 TGTCTTATCTGAGACCTGCGCGCTTGTGAGCTGTGCTTCAACCTCCACCAAACTATC 2417
Qy 2101 GTGAGTCAATTAATGATGCTATCACTGCTTCACTGCTTCAAAATACATGTCGATGG 2160
Db 2418 GTGAGTGTGAATACCTGTATAGGTATAGGTGTGGAGTGTCTCTTGTGAATCAATGG 2477
Qy 2161 GAGTGGTATTAATCTTATTTCTGCTCTTAAGCGAGACCAAGGTTTGGCTGTGATGG 2220
Db 2478 GAGTATATCTGTTGCTTCTTCTTGTGCGGAGCGCGCTGTGTGCTGTGCTGTG 2537

QY 2221 ATGTCATCTTGTGGCCAGGCGGAAGCAGCAGTGAAGAACTGTGATCTTTCAGAGCT 2280
DB 2538 ATGATGCTGCTGATTAACCCAGGCTGAGGCCACTTAAAGAACTGTGTGCTCTAAATGCG 2597
QY 2281 GCGAGCGCAGCTAGCTGCATAGGCTTCTGTAATTTTGTGATCTTTTTCGTGGCTGTGG 2340
DB 2598 GCGTGTGTGGCCGAGGCGCATAGGCTTCTCTCTCCCTCGTGTCTTCTGTGGCCGCTGG 2657
QY 2341 TACATCAAGGCTCGGGTAGTCCCTTAGCTACCTATCCCTCACTGGCCGTGTGCTCTT 2400
DB 2658 TACATCAAGGCGAGCTGTGCTCTGGGCGGCGATATGCTCTTAAAGCGATAGGCGGTG 2717
QY 2401 AGCCTACTGCTCTAGCATTTGCCCAACAGGCTTATGCTTATGACGATCTGTGATAGC 2460
DB 2718 CTCTGTGCTTGTGCGCTTACCAACAGACTTATGCGACCGAAGATGGCTGCA 2777
QY 2461 CAGATAGGAGCGGCTGTGCTGTATATGATCACTCTTTTACTCTACCCCCGGTATAG 2520
DB 2778 TCGTGGGAGGCGCGGTTTTTGTAGTCTGTACTCTTGTACCTTGTACCATACTATAG 2837
QY 2521 ACCCTTCTCAGCGGTTTTTGTGGTGTGTGCTATCTTGACCCCTGGGGGAAGCTATG 2580
DB 2838 GTGTTCTCGCTAGCTCATATGGGTATACAAATATTTTATCACAGAGCGGAGCGCAC 2897
QY 2581 GTCCAGAGTGGGACACACTATGCAAGTGGCGGCTGGATGATCATATGAGGCC 2640
DB 2898 TTGCAAGTGTGGTCCCTCTCTGATGTTGGGGAAGCGCGATGCTATCTCTT 2957
QY 2641 GTCCCATATTTCTAACCCAGGTGTGTGTTGATCAATACCAAGTGGCTTGTGGGCTT 2700
DB 2958 ACATGCGGCGTCCATCCAGAGCTAATCTTTGACATCACAAACTCTGCTGCATCTC 3017
QY 2701 GGGCTGTGTTACCTCTTAAAGGTGCTTGAAGCGGCGGTACTCTGTACAGGCTCAC 2760
DB 3018 GGTTCGCTCATGTGTCTCAGGCTGGCATTAAGAGTGGCTGATCTTGTAGCGGCTCAG 3077
QY 2761 GCTCTACTGAGAGTGTGACCATGAGCATCTCGGCGGCGGAGGTAGTGCAGATG 2820
DB 3078 GGGCTCATCCGTGATGATGTTAGTGGAAAGTCCGTGGAGGCACTAATGTCAAATG 3137
QY 2821 GCGCTACTAGCCCTTGGCAGGTGACTGGCACTTATCATTTATGACACTGACCCCTATG 2880
DB 3138 GCCTTCATGAAGCTGGCCGCTGACAGATAGTACGTATATGACATCTTACTCCACTG 3197
QY 2881 TCGATTTGGGCTGCTAGTGGCTGGGACCTGGGCGGTGGCGCTTGAAGCTATCATCTTC 2940
DB 3198 CGGATTTGGGCCACGCGGCGCTTAGAGACTTGGGTGGCAGTAGAGCCGTGTCTTC 3257
QY 2941 AGTCCGATGAGAGAAAGTCAATTTGTGGGAGCGGAGCAGCTGTGTTGGGAGACTT 3000
DB 3258 TCTGACATGAGAGCTAACTCATCACTGGGGGGAGACACGGGGGTGTGGGAGATC 3317
QY 3001 TTACACGAGCTTCCCGTGTCCGCCGACTTGTGTGGAGGTCTCTTGGCCGACTGTAT 3060
DB 3318 ATCTCGGCTTACACAGTCTCCGCCGAAAGGAGATATCTTATAGGACCCGCCGAT 3377
QY 3061 GGCTATACCTCAAGGGGTGAGTCTTGCCGCCCATCACTGTCTAGCCGACAGACGA 3120
DB 3378 AGTTTGGAGAGCAGGGGTGGCGGCTCTTGCGCTTACAGGCTATTTCCAAACAAACG 3437
QY 3121 CGTGGCCCTTTTGGCACCATAGTGTGAGATGACGGGGCGCGACAGAGACGAACAGCT 3180
DB 3438 CGGGGCGCTGTGGGTGTATATATCACTAGCTTACAGAGTGGGACAAAGATCGAT 3497
QY 3181 GGGGAATTCAGGTCTGTGCAAGTCACTAGTCTTCTGGAATCATCTATCTCGGG 3240
DB 3498 GGGGAGGTTCAAGTGTCTTCCACCGCAATCTTTCGTGGGACCTGTGCTAAATGGC 3557
QY 3241 GTTTTGTGAGCTGTCTACCATGAGCTGGCAACAAGATCTGGCGGCTCACGGGGTCCG 3300
DB 3558 GTGTGTGAGCCGTCTACCATGTGTGCGGCTCGAAGACCTGGCGCGCCGAAAGGTCTCA 3617

QY 3301 GTCCAGAGATGATCTCAAGTGTGAGGGGAGCTTATAGTGGGTGGCCAGCCCTCTGGG 3360
DB 3618 ATACCCCAATGTAAACCAATGTAGACAGAGACTGTGGGCTGGCGGGGCCCTCCCGGG 3677
QY 3361 ACTAAATCTTTGAGCGGTGACAGTGTGAGCGGTGACCTGTACTGTGTACGCGGAAC 3420
DB 3678 GCGCGCTTCATGACACCTGTGACCTGTGGAGCTCGGACCTTATCTTGTGTACAGAGCAT 3737
QY 3421 GCTGATGTATCCCGGCTGAAAGCGGGGAGCAAAAGGGGAGCGCTACTCTCCCGAGA 3480
DB 3738 GCTGATGTGTGTTCCGATGTGCGCGCGGCGAGACAGAGGGGAGAGCTGTCTTCCCGAG 3797
QY 3481 CCTCTTTCACCTTGAAGGGGTCTTCAAGAGCGCCGCTGTATGCCCCAGGGGACAGCT 3540
DB 3798 CCCATCTCTTACTGAAGGGCTCTCGGGTGAACACTGCTTGTCCCTTGGGGGACGTT 3857
QY 3541 GTTGAAGTCTTCCGGGACCTGTGTGCTCTCGGGGTGTGGCTTAAATGCTCAATTTATC 3600
DB 3858 GTAGGATCTTCCGGGCTGTGTGACCCGGGGGGTGGCAAGGCGGTGACCTTCATA 3917
QY 3601 CCGGTGAGACACTGACATGCTCAAGCGGTCCCGCACCTTTATGTACAAAGACACCA 3660
DB 3918 CCGTTGATGCTATGGAACACTACATGCGGTCTCGGCTCTTACAGACAACTCATCTCC 3977
QY 3661 CTTGCTGTGCCGACACTATCAAGTGGGTACTTGATGATGCCCGACTGGCAGTGAAG 3720
DB 3978 CCGGCGGTACCGCAAACTTCCAAATGTGCAATTTACAGCTTCCACTGGCAGCGCAAG 4037
QY 3721 AGCACCAAGTCTCTGTGCTATGTGCTCTCAGGGGTATTAAGTGTAGTGTATGCC 3780
DB 4038 AGCACCAAGTGTGCGGCTCATATGACGCCAAGGGTAAAGGTGTCTGTCTTAAACCG 4097
QY 3781 TCAGTGTGCGCAACCTCGGGGTTTGGGGGTACTTGTCTAAGGCAATGGCATCAATCC 3840
DB 4098 TCCGTTGCGGCACTATGGGCTTTGAGGCTATATGTCCAAAGGACATGGATCGAGCT 4157
QY 3841 AACATTAGACTGTAGTCAAGACTGTGACGACCGGGGCGCCCATCAAGTACTCCACAT 3900
DB 4158 AACATCAAACTGGGGTAAAGAACATCAACACGGGCGGCCCATCAAGTACTCCACT 4217
QY 3901 GGCMAATTCCTGCGGCTATGGGGCTGTGGGGCGGCTTACGACATCATATATGTGAT 3960
DB 4218 TGCMAATTCCTTGGCCGAGGTGATGCTCGGGGGCGCTATGACATCAATATATGTGAT 4277
QY 3961 GAATGCAATGCGGTGACTCTACCAACATCTTGTGATCGGAACAGTCTTGTCAAGA 4020
DB 4278 GAATGCACTCACTGACTGACTGACTCACTATCTTGGGATGGGACATCTCTGATCGGCA 4337
QY 4021 GAGACAGCTGGGTGAGACTTAATCTGTGTGACTTACAGTACGCGCCCTGGGTCAGTGAC 4080
DB 4338 GAGACGCGTGGAGCGGGCTCGTGTGTGCGCACCGCCACTCGGGATGATCACCC 4397
QY 4081 ACCCCCCCAACCAATAGAGAGTGGCTTGGGCAAGAGGGCGAGATCCCTTCTAT 4140
DB 4398 GTGCCACACCCCAACATCAAGAGTGGCCCTGTTCACAACTGGAAGATTTCCCTTCTAT 4457
QY 4141 GGGAGGCGCATTTCCCTGTCTTACATCAAGGAGAGAAAGATCTGTCTTCCCATTTCA 4200
DB 4458 GGGAAAGCATTCCCATTTGAGGCTATCAAGGGGGAGAAAGCATTTCTTCTGTGATTC 4517
QY 4201 AAGAAAAAGTGAAGAGCTCGGGCGGCCCTTTCGGGATATGGGCTTGAATCTCAGTGCA 4260
DB 4518 AAGAAAGATGTACAGAGCTCGCGCAAAAGCTGACAGGCTCGGATCTCAATGTGTAGCG 4577
QY 4261 TACTACAGAGGTTGAGAGCTCTCCGTAAATACCACTCAGGGAGACGTAGTGTGTGCGCC 4320
DB 4578 TATTACCGGGGTCTGATGTGTGCTGTATACCACTAGCGGAAGCGTGTGTGTGCGGCA 4637
QY 4321 ACCGAGCGCTCATGACAGGATATCTGGGAGCTTTGACCTCGGTATGCACTGCAAGTA 4380
DB 4638 ACAGAGCTTAAATGACGGGTTTACCGGAGACTTTGACCTCAAGTATGACTGCAACACA 4697
QY 4381 GCGGTCACTCAAGTTGTAGCTTCAAGTTAGACCCCACTTCACTAATCAACACAGATT 4440

4698 TGTGTCAACCCAGACAGTGCATTTTCAGCTTGGATCCCACTTACCATTTAGAGACGACAAG 4757
4441 GTCCCTTAAGACGCTGTCTCACTGTAAGCCAGCGCGGGGTGCACGGGTAGGGGAAGACTG 4500
4758 CTGCCCCAAGACCGGGGTGTGCGCTGCCAGCGCGGAGGTAGGAACTGGCAGGGGAGAGAT 4817
4501 GGCAATTAATAGATATGTTTCCACTGTGTAGGAGAGCCCTCAGGAATGTTTGA CAGTGTAGTG 4560
4818 GGCATCTACAGGTTTGTGTGATCTCCAGAGAAAGGCCCTCAGGCATGTTCAGTCTCGGTG 4877
4561 CTCTGTAGTGTCTACGACGAGGGGGCGCATGTGTAGCTCACACCATCGAGACCAACC 4620
4878 CTGTGTAGTGTCTATGACGAGGGCTGCGCTTGGTATAGAGCTCAGCCCGGTGAGACCTCG 4937
4621 GTACAGCTCAGGGCGGTATTTTCAACAGCCCGGTTTGGCTGTGTGTCCAAAGCATCTTGA 4680
4938 GTTAGGTGTGGGCTTACCTAAATACACAGGGTTGCCGTGTGCCAGAGCAACTTGTAG 4997
4681 TTTTGGAGGCGAGTTTTCACCGGCTCAACAACATAGATGCCCACTTCTTCCCAACA 4740
4998 TTTCTGGAGAGCGTCTTCAAGGCTCAACCACTAATGATGCCCACTTCTGTCCCAAGCC 5057
4741 AAGCAATCGGGGGAATTTTCGATTAATTAACAGCTTACAGGCTTACAGTGTGCGCTAGG 4800
5058 AAACAGGACGAGACAACTCCCTTACCTGTGTAGCATACCAAGCACAAGTGTGCCAGG 5117
4801 GCCAAAGCCCCCCCCCGTCTGTGGAGCATGTGTGAAGTGTGTGACTCGACTCAGACCC 4860
5118 GCTCAGGCTCACCCTCATGTGTGAGCAAAATGTGAAAGTGTCTCATACGGCTTAAAGCCC 5177
4861 ACACTGTGGGGCCCAACCTCTCCTGTACCGCTTGGGCTCTGTATCAACAGAGTACC 4920
5178 ACCTGTGATGGCCAAAGCCCCGTGTGTACAGGCTAGAGGCCGTTCAAAATGAGTACT 5237
4921 CTCAACATCCCGTGAAGAAATATCATGCCACCTGTACATGCAAGCCGACTTGTGAGTATG 4980
5238 CTCAACACCCCAATACCAATATCATATGAGCATGATGTCCGTGACCTGTGAGTGTCT 5297
4981 ACCAGCAATGGGTCTTTGGCAGGGGAGTCTTGGCGGCGCTCGCCGCTTATTTGCTGGCG 5040
5298 ACTTAGACCTGGGTGTGTAGTAGGCGGAGTCTTGGCGCTCTGGCGGCTGTCTGCTAGC 5357
5041 ACCGGGTGTGTTCATCATCGCGCCGCTTGCACATTTAACAGAGAGCCGCTGTTCGCCCG 5100
5358 ACAGGAGCGGTGTCTGTGTGGGAGGATATCTTGTCCGGAGGCCAGCTGTATTTCC 5417
5101 GACAAGAGGTCTCTTATGAGGCTTTTGTATGATGAGTGAAGATGTGCTTAGGCGGCT 5160
5418 GACAGGGAATCTCTTACAGAGTTCGATGAGATGGAAGATGTGTCTTACACCTCCCT 5477
5161 CTGATTGAAAGAGGGGACCGGATAGCCGAGATGCTGAAGTCCAAATCCAAAGGCTTATTTG 5220
5478 TACATGAGAGCAAGAAATGCACTCGCCGAGCAATTTCAACAGAAAGGCGCTCGGATTTGCTG 5537
5221 CACCAAGCTTCCAAAGCTCAAGCTCAAGCATACAAACCACTGTGAGGCTTATAGGCCCAAG 5280
5538 CAAACAGCCACCAAGCAGAGGAGGCTGTCTCCGCTGTGAGTCAAGTGGCAGACC 5597
5281 GTAGAACATTTCTGGGCAAAACATGTGAACTTCTTAGCGCATCCAAATCCCTGCA 5340
5598 CTTGAGGTCTTCTGGGGAAACATGTGAACTTCTACGCGGGAATACGATCTTTGGCA 5557
5341 GGACTTACCACTGCCAGGAAACCTGTGAGTGTTCATGATGGCTTCACTGTCGCC 5400
5658 GGGCTTATCCACTCTGCTGTGAAACCCCGCATAGCATATGATGTGCTTTTAAAGCTCTCT 5717
5401 CTCACAGTCCGCTGTCACAAAGACCACTATCTTCTCAACATTTTGGGGGGCTGGCTA 5460
5718 ATACACAGCCCGCTCACACCAAAATACCTCTCTGTTTAACTTTGGGGGGAATGGGTG 5777
5461 GCATCCCAATTCACCAACCGCGGGGCCATGTGCTTCTGTGTCACTGTGCTTAGTGGGA 5520

5778 GCTGCCCACTCGCTCCCGCCAGCGCTGCTTCGAGCTTTCGTGGGCGCGCGGACTTGCCTG 5837
5521 GCTGCCGTAGGCAATATAGCTTAAAGTATAGTGTATGTGACATCTCTGGCAGGCTATGT 5580
5838 GCGGCGTGTGGCAGCATATAGTCTCGGGAAAGTCTTGTGACATTTCTGGCGGGCTATGGG 5897
5581 GCGGGAATTTCCGGGGGCTCTGTGCAATCAAGTCAATGTCTGGGCGAAGACCTTCATG 5640
5898 GCGGGGTGTGCTGCGCACCTGTGCTTTTAAAGTCAATGAGCGGACGAATGCTTCCACT 5957
5641 GAGATGTGTCTCAACTTGTGCTGCTGGAATTTCTGTCTCGGGGTGCTTGTATGTGGAGTC 5700
5958 GAGATGTGTATTTACTTCCCTGCGCATCTTCTCTGCGCGCTGTGTGTGCGGGTCT 6017
5701 ATCTGGCGGCGCAATTCGTGCGCCACACATGTGGAGACCGGGGGAAGCGCGCTTCAATG 5760
6018 GTGTGGCACAATATCTGCTGTGACATGTGGGCGCGGGAAGGGGCTGTGCACTGTAGT 6077
5761 AATAGACTATTTGCTTGTCCAGAGGAATTCAGTGCSCCCCAACCACTACGTAGC 5820
6078 AACCGGTGATTAAGCTTCCGCTTGCGGGGTAAACAGCTCTCCCAAGCATATGTGCC 6137
5821 GAGTCGATGTGTCGACGCGTGTACCCAACTATCTTGTGCTTCCCTTACATTAACAGCTG 5880
6138 GAGAGCGACGCGCGCGGCTGTATTAAGATCTCTCCAGCTTACATCACTACATGTTG 6197
5881 CTCAAGAACTCCACAACTGTGATTAATGAGAGATGTGCCCATCCCATGTGGCGGCTGTGG 5940
6198 CTAAAGAGGCTTATCATGTGTGATTAATGAGAGATGTGCCACGCTTGTTCGGCTGTGG 6257
5941 CTCGCGAGTGTGTGGGACTGGGTTTGGACATCTCTAACAAGCTTTAAATAATGAGTACC 6000
6258 CTAAAGATGTGTGGGACTGTGATATGACAGGCTGTGTGATACCTTCAAGACTTGTGGCTCAG 6317
6001 TCCAAATTAATCCCAAGATGCGCGGCTCTCCCTTGTCTCTGTCAAAAGGGGTACAG 6060
6318 TCCAGGCTCTGCGCGGTTACCGGAGCTCTTCTCTGTATGACCAACCGGGGTACAG 6377
6061 GGGGTGTGGCGCGGCACTGTGATCATGACACACAGGTGTCTTGGCGGCGCAATTTCT 6120
6378 GGAAGTGTGGGGGAGATGATCATATGCAAAACACCTGCCCATGTGAGACAGATAC 6437
6121 GGCATGTCCGCTTGTGGGCTCATAGAAATACAGGGGCTTAAAGCTGTATGAATATCTG 6180
6438 GGAATGTCAAAATATGCTTCTCATAGGATTTTGGCCAAAACTGTGACAAACCTG 6497
6181 CAGGGACCTTCTCTATCAATTTGTATACAGAGGCGCAGTGTGCGGAACCGCGCCA 6240
6498 CATGAACATTTCCCATCAACGATACACAGGGGCGCTGACGCGCTCCCAAGGCGCG 6557
6241 AACTTTAAGTGTGCGCATGTGAGGGTGTGGGCTCAGAGTATGCGGAGGTGAGCGAGAC 6300
6558 AACTTATTCAGGGCGGTGTGTGGCGGTGTGCTGAGAGATGATGTGAGAGTTTACGCGGTG 6617
6301 GGGTCAATACCACTACATTAACAGGACTCACCACTGATTAATCTTAAAGTCCCTCCAACTA 6360
6618 GGGGAATTTCACTACGTGAGGGGATGACATGACATGACAGGTGAATGTCCATGTCCAGGTT 6677
6421 CCGAAGCGTTTTTCCGGGATGAGTGTCTGTTGTGGGTGTGGGTTTAACTTGTGTCTG 6480
6738 TGCAAACCTCTCTTACGAGAGAGGTGTGATTTCCAGGTGTGGGCTCAACAGTACTGTGTC 6797
6481 GGGTCCCAAGTCTCTTGTGCAACCTTGAACCGGACAGACAGATATTTAGTGTCACTATACA 6540
6798 GGGTCAAGCTCCCATGTGAGCCCGGAACCGGATGTGTGAGTGTCTCACTGTCTAC 6857
6541 GATCATCTTCAATACAGGCGAGACTGACGCGGCTTTTACGCGGGGGTCAACCCCA 6600
6858 GACCCCTTCATATTTACAGCAGAGAGCGGCAAGCGTATGAGCTGTGCCAGGGGGTCTCCCCC 6917

QY	6601	TCGAGGGAAGACTCTCTGGGCGAGCGAGTATCGGACACATCGCTGGCAGACCTCGACC	6660
Db	6318	TCCTTTGGCGACACTTTACAGCTTAGCGACGTTGTCCTCGCCTTCTTTGAAAGCGACATGACT	6977
QY	6661	ACCCACGGCCAAAGCCTATGATGTGAGACATGCTGATGTAACCT-----GTTCC	6708
Db	6978	ACCCATCATGACTCCCGGAGCGCTGACCTCATGAGGCGAACCTCTGTGGCGGCAAGAG	7037
QY	6709	ATGGGGGGCCGATGTGACTCGGATAGAGTCTGGGATCCAAAGTGTGCTTCTGGACTCTCTC	6768
Db	7038	ATGGCGCGGAAACATCACCCCGTGTGAGTCAAAAAATAAAGGTGTATCTGGACTCTTTTC	7097
QY	6769	GACCCATGTGTGSAAGAAAGGAGGACCTTAGGCGCTTGATACATGACAAATATAGTCTC	6828
Db	7098	GATCCGATTCGGGCGGTGTGAGATGAGAGGAAAAATTCGTCGCGCGAGATCTTGGA	7157
QY	6829	CCCAAGAAAGAGTTCCCAACAGCTTTACCGGCGTGGGCGACGGCCTGTATTACAAACCA	6888
Db	7158	AAACCCAGGAAGTTCCCGCCGACCGTGGCCATATGGGCGACCGCGAATTTACAACTCTCA	7217
QY	6889	CTTGTGGAAATGTGTGAAAAAGGCCAGATTACCAACCGGCACCTGTTGGCGGCTGTGCTCTC	6948
Db	7218	CTGTAGAGTCTCTGAAAGAACCCGAGACTACCTCCCCGGTGTATCAAGGGTGTGCCCTTTG	7277
QY	6949	CCTCCTCCTTAGAAAAACCCCGACGCGTCCCCCAAGAGAGCGCGGACAGTGGGCTPAAGT	7008
Db	7278	CCATCTACCAAGGCCCCCAATACACTCTCAAGAGAAAGAGACGTTGTCTTGACA	7337
QY	7009	GAGGACTTCATAGAGAGATGACCCTTCAACAGCTGGCCATTAAATGCTCTTTGGCACGCCCC	7068
Db	7338	GAGTCCACCGTGTCTTCTGCTTGCGGAGGCTCGCTACTAAGACCTTTTGGCAGCTCCGGG	7397
QY	7069	CCAAAGCGGCGATTACAGGCTTTCCACGGGGGCGGGCGTGGCCGATTCCGGCAGTCAAGC	7128
Db	7398	TCGTGCGCGCTGTGACAGCGG---CACGCGCACTGGCCCTCCCGA-----TCAAGGCC	7445
QY	7129	CCTCCTGATGAGTTGGGCCCTTTCCGAGACAAGTTCCATCTTCCATGACCCCCCTGGAG	7188
Db	7446	TCGCGACGCGCGCAGCAAAAGATCCGACGTTGATGTGTACTCTCCATGCCCCCCTCGAG	7505
QY	7189	GGGGAGCTTGGAGATCCAGACCTGGAGGCTGAGCAGTATGAGCCCAACCCCCCCCGAG	7248
Db	7506	GGAGAGCCAGGGGAGCCCGGACT-----	7528
QY	7249	GGGGGGGTGGCAGCTCCCGGCTCGGACTCGGGGTCTGTCTACTTGTCCGAGAGGAC	7308
Db	7529	-----CAGGACCGGGCTTGTGTTACCGGTGAGGGGAAAGCTGTG	7568
QY	7309	GACTCCGTCGTGTGCTGCTCCATGTCATATCTCGGACCGGGGGCTCTAATAACTCTTTGT	7368
Db	7569	GAGGACGTGTCTGTCTGCTCATATGTCCTATATAGACAGGTGCTTGTATCACGCCATATC	7628
QY	7369	AGTCCCGAAGAGAGAAGTTAACCCCTTGAGCAACTCCCTGTGGTCGATATATCAC	7428
Db	7629	GCTGCGAGGAGAGAGCAAGTTGCCATCAATCCGTTGAGCAACTCTTTGCTCGTCAACAC	7688
QY	7429	AACAAGGTGTACTGTATCCACAACAAGAGCGCTTCATCTAAGGGCTTAAAAAGTAACTTTT	7488
Db	7689	AGTATGTGTCTACTCCACAACATCTCGACGCCAAGTCTGCGGCAAGAAAGTCACTTTT	7748
QY	7489	GATAGAGATGCAAGTCTCGACTCCACTATACGACTCAGCTTAAAGGACATTAAAGTACG	7548
Db	7749	GACGACCTGCAAGTCTTGAGCAACACTACCGGAGCGTGTCTCAAGAGATGAGGCGAAG	7808
QY	7549	GCTTCCAAAGTCAACCGCAAGGCTCCTCACTACATGAGAGAGGCTTGCAGTTAAACCCACCC	7608
Db	7809	GCGTCCACAGTTAAAGGCTAAGGTTCTATATAGAGAGAGGCGCTGAAACTACGCGCCCA	7868
QY	7609	CATTCTGCAAGATCTTAAATATGTGGTTTGGGCTTAAGAGAGTCCGACGTTGTCCGGAGG	7668
Db	7869	CATTTCGCAAAATCCAAATTTGGCTACCGGGGGAAGGAGCGTCCGAGGCTATTCAGCAGG	7928

QY	7669	GCGGTTAACCAATCAAGTCGCGGTGGAGAGACCTCGTGGAGAGCTCAGAAAACCCAAATT	7728
Db	7929	GCCTTCAACCAATCCGCTCCGTGTGGAGAGACTTGTGGAAAGACTGAAAACCACTT	7988
QY	7729	CCCAACAACATATAGCCAAAATAGAGTGTCTTGCCTGGACACCCCAAGGCGGCAAG	7788
Db	7989	GATACCAACATCATGAGCAAAAAAGAGAGTTTCTGCTCCAAACAGAAAAGGAGCGCG	8048
QY	7789	AAAGACACTGGCCCTTATTCGTTTACCTGACTCGGCGCTCAGGGCTTCGAGAAAGATGGCC	7848
Db	8049	AAGCAGCTGGCTTATTCGATTCCTCCAGACTGGGGGTACGTGATTCGAAAGATGGCC	8108
QY	7849	CTTTATGACATATACAAAAAATCTTCAGGCGCGTGTGATGGGCGCTTCTTATGATTCAG	7908
Db	8109	CTTTACGACGCGTCCACCCCTTCCTCAGGCGCGTGTATGGGCCCTCATACGAGATTCCAG	8168
QY	7909	TATTTCCCCGCTCAGCGGGTATGAGTTTCTCTTAAAGCATTGGCGGCAAAAAGAGACCTT	7968
Db	8169	TACTCTCCTGGGACGCGGTGAGTTCTTCGTGTAAATCTGGAAATCAAGAAATGCGCT	8228
QY	7969	ATGGGTTTTCTGATGTGATACCCGATGCTTTGACTCAACCGTCACTGAGAGACATCAGG	8028
Db	8229	ATGGGCTTCTCATATGACACCCCGTGTCTTGTGACCTCAACGGTCACTGAAATGACATCCGT	8288
QY	8029	ACTGAGAGTTCATATATCGGGCTGCTCCTTGCCTGGAGAGGCGCAACTGCCATACAC	8088
Db	8289	ACTAGGAATCAATTTTACCAATGTTGTGACTGTGCCCCCAAGCAGCAGCGCCATATAGG	8348
QY	8089	TCGCTAATCTGAGACATTTACGTGGAGGCGCTATGTTTCAACAGCAAGGCGCAAACTGTC	8148
Db	8349	TCGCTCACAGACGCGCTTATATGTGGGGGTCCCTGTGACTCAACGGTCACTGAAATGCAACTTCG	8408
QY	8149	GGGTACAGCGCTTGCACGCGCCAGCGGGGTGTCCACCTAGCATGGGAAACACCATGACA	8208
Db	8409	GTTATATCGCGGTGCTCCGCGCAAGTGGCGGTGTGACACTAGCTGCGGCAACACCTCACA	8468
QY	8209	TGCTACGTGAAAAGCTTAGCGGCTTTGTAAAGCTGCAGAGATATATCGGCGCCACAAATCCTG	8268
Db	8469	TGTTACTTGAAGGCGCTTACGTGCGCTGTGTCGAGCTGCAAAAGCTCCAGGACTGCATATCTC	8528
QY	8269	GTATGCGCGCATGTACTGTGTTCATCTCAGAAAACCGAGGGAGCCGAGAGAGACGAGCGG	8328
Db	8529	GTGAACGAGACGACTTGTTCGTTATCTGTGAGAGTGCGGGAACCCAGAGAGATGCGGCG	8588
QY	8329	AACTGAGAGCTTCAACGAGGCTATGACAGAGTATTTCTGCCCCCTCTGTGTACCCCCC	8388
Db	8589	GCCCTACGAGCTTCAACGAGGCTATGACTAGATTTCCGCCCCCGGGAGACCGCCCG	8648
QY	8389	AGACCGAGTATGATCTGAGAGCTATTAACATCTTGCTCCCTCAATGTGTCTGTGGGCGCTG	8448
Db	8649	CAACCGAATACGACTTGGAGCTATTAACGATGCTCTCTCCATGTGTGCGTCCGAC	8708
QY	8449	GGCCCAACAGGCGCGCGATACTACTCTGACAGAGACCTTACACTCCAATGCGCCCG	8508
Db	8709	GATGATATCCGCAAAAAGGTGTACTACTCCCGGTGACCCCAACACCCCCCTCTGCACGG	8768
QY	8509	GCTGCTGGGAAAACAGTTTGAACACTCCCGTCAATTCAATGAGCTGGGAAAACATCATCAG	8568
Db	8769	GCTGTGTGGAGACAGTTTGAACAATCCAGTCACTCTCGGCTAAGGCAATATATCATG	8828
QY	8569	TACGCCCCGACATATGGGCTCGCATGGTCTCTGTATGACACACTTTCTTCCATTTCTATG	8628
Db	8829	TATGGCCCAACCTATGGCGAGATGATTTGTATGACTCATTTCTTCTGTATCTTCTTA	8888
QY	8629	GCTCAAGACACGCTGAGACAGAACCTCAACTTTGATGATGTAAGGAGCGGTATCTCGGTG	8688
Db	8889	GCTCAGGAGCAATGTGAAAAAGCCCTGATTTGTACGATCTTACGGGCGCTGTACTTCATTT	8948
QY	8689	AGTCCCTTGAACCTCCAGCTATTAATTGAAGTTTACATGGGCTTGAACGCTTTTCTCTG	8748
Db	8949	GAGCCACTTGAACCTACCTCAATCATTTGAACAGATCCATGATCTTTAGCGCATTTTCACTC	9008
QY	8749	CACATATACATCTCCCAAGAACTGACACGAGTGGCTTCAGCCCTCAGAAAATTGGGCGG	8808

Db CACGACTCTCTCAGGTGAGATCAATAGGCTGCTTCATGCTCAGGAACTTGAGGTA 9068
Qy 8809 CCACCTCTCAGAGGTGAGGAGAGCCGCGACGTGCTCAGGAGCTCCCTCATCTCCGT 8868
Db 9069 CCGCTTTTGAGAGTCTGAGAGCATCTGGGCGCCAGAGTCTCCGCTTAAGTCTCCAG 9128
Qy 8869 GGGGAGAGAGCGGCGTTTGCGGTGATATCTCTTCAATTGGGCGGTGAAGACCAAGCTC 8928
Db 9129 GGGGAGAGAGGTGCTGCTCAGCTTGGCGAGTACTCTTCACTGGGAGTGAAGACCAAGCTT 9188
Qy 8929 AAACCTCACTCCATTGCGGAGAGCGGCTCTGATTTATTCACGTGTTACCGTGGC 8988
Db 9189 AAACCTCACTCCAAATCCGGGCTGGGTCCAGCTAGACTTTGTCGGGTGTTGCTGCTGT 9248
Qy 8889 GCCGCGCGGAGGAGCATTTATCAGAGGTGCTGCGGTCCGAGCCCGCTTATGCTCTTT 9048
Db 9249 TACAAACGGGAGAGCATATATACAGGCTGTCTGTGCTCCGACCCCGTTGGTTCATGTTG 9308
Qy 9049 GGCTACTCTCACTTTTGTAGGGGTAGGCTTTTCTCACTCCCGGCTCG 9099
Db 9309 TGCTACTCTCACTTTGTAGGGGTAGGCACTTACTCTCCCAACCG 9359

RESULT 14

ADP88596 standard; DNA, 9413 BP.

AC ADP88596;

DT 26-FEB-2004 (first entry)

DE Hepatitis C virus NS3 gene, SEQ ID No 5.

KW liver cancer; hepatitis-C virus; NS3 gene; carcinogenicity; ds; gene.

OS Hepatitis C virus.

PN JP2003210181-A.

PD 29-JUL-2003.

PF 30-MAY-2002; 2002JP-00158335.

PR 16-NOV-2001; 2001JP-00352443.

PX (SHIN-) ZH SHINSANGYO SOZO KENKYU KIKO.

PA WPI; 2003-819836/77.

DR P-PADB; ADP88597.

PT Diagnosing liver cancer, involves amplifying amino terminal region of hepatitis-C virus gene using predetermined primer and determining hepatitis-C virus in base sequence of obtained DNA fragment.

PS Disclosure; SEQ ID NO 5; 36pp: Japanese.

XX The invention relates to the novel testing method for diagnosing liver cancer. The novel method comprises amplifying the amino terminal region of a hepatitis-C virus NS3 gene using a predetermined primer and determining the hepatitis-C virus in a base sequence of the obtained DNA fragment. The novel testing method is useful for diagnosing liver cancer and also used in a gene amplification technique, a clinical laboratory test reagent, a polymerase chain reaction, a base sequence analysis and genetic engineering. The method enables the detection of a hepatitis-C virus having high carcinogenicity with high specificity. This CC polynucleotide sequence represents the hepatitis-C virus NS3 gene of the invention.

SQ Sequence 9413 BP; 1886 A; 2816 C; 2694 G; 2017 T; 0 U; 0 Other;

Query Match

Best Local Similarity 45.5%; Score 4141.8; DB 10; Length 9413;

66.6%; Pred. No. 0;

Matches 6066; Conservative 0; Mismatches 2952; Indels 93; Gaps 6;

Qy 1 ATGAGCAAAATCTTAACCTCAAGAAAACCAAAAGAAACCAACCGTCCGCCCA 60
Db 330 ATGAGCAAAATCTTAACCTCAAGAAAACCAAAAGAAACCAACCGTCCGCCCA 389
Qy 61 GACGTAAATTTCCGGGCGGCGGACAGATGTTGGCGAGATATCTTTGGCGGAGG 120
Db 390 GACGTAAATTTCCGGGCGGCGGAGTCAATGTTGGAGATTTACCTGTTGGCGGAGG 449
Qy 121 GGGCCAGGTTGGGTGTGCGCGCAAGAAAGATTGAGAGCGGTCCAGCCAGTGA 180
Db 450 GGGCCAGGTTGGGTGTGCGCGCAAGAAAGATTGAGAGCGGTCCAGCCAGTGA 509
Qy 181 AGGCGCAGGCAATCCCTTAAGATCCGCGCTCACTGGGCAAAATCTGGGAAAACAGGA 240
Db 510 AGGCGCAGGCAATCCCTTAAGATCCGCGCTCACTGGGCAAAATCTGGGAAAACAGGA 569
Qy 241 TACCCCTGAGCCCTATACGGGAAATGAGGAGCTCGGCTGGGCAAGATGGCTCCGCC 300
Db 570 TACCCCTGAGCCCTATACGGGAAATGAGGAGCTCGGCTGGGCAAGATGGCTCCGCC 629
Qy 301 CGAGTTCCGCTCTCTTGGGAGCCCAATGACCCCGGATAGTGGCCAACTGAGGT 360
Db 630 CGTGGCTCTCGGCTTATGTTGGGAGCCCAAGACCCCGGATAGTGGCTTATGAGGT 689
Qy 361 AAGTCAATGATACCTTAACGTGCGGCTTGGGCACTTATGGGATATCCCTGTCGTG 420
Db 690 AAGTCAATGATACCTTAACGTGCGGCTTGGGCACTTATGGGATATCCCTGTCGTG 749
Qy 421 GGGCCCCGCTCGGCGGCGTGCAGAGCTCTCGGCAATGAGGAGTCTGAGAGAG 480
Db 750 GGGCCCCGCTCGGCGGCGTGCAGAGCTCTCGGCAATGAGGAGTCTGAGAGAG 809
Qy 481 GGGGTTAATTTTGCAACAGGAACTTACCGGTTGCTCTTTCTATCTTCTGTCGCC 540
Db 810 GGGGTTAATTTTGCAACAGGAACTTACCGGTTGCTCTTTCTATCTTCTGTCGCC 869
Qy 541 CTGCTGCTGATACCAACCCCGGCTCTCGGCTGCGGAAATGAGGAAATCAGTACCGGC 600
Db 870 TTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 929
Qy 601 TACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 930 TACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 989
Qy 661 CTCACGCTCCCGGAGTCTCCGTCGAGAGAAAGTGGAGATGATGATGATGATGATGATGAT 720
Db 990 ATGACACACCCCGGAGTCTCCGTCGAGAGAAAGTGGAGATGATGATGATGATGATGATGAT 1049
Qy 721 CCGGTCTACCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 1050 CCGGTCTACCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1109
Qy 781 CACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 1110 CACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1169
Qy 841 TGGGTTGGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 1170 TGGGTTGGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1229
Qy 901 GTCCAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 1230 GTCCAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1289
Qy 961 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 1290 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1349
Qy 1021 CCGAGGTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Db 1350 CCGAGGTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1409

Oy	1061	TACTTCTCTAAGCAGGAGAGCGTGGGCGAAATCGTGTGACATCCCTTCTGTTGGCGCGCGG	1140
Db	1410	TACTATTCTCATGATGATGGGAACTGGGCTTAAGGCTTTGATTTGATGCTACTTTTGTCTGGC	1469
Oy	1141	GTGACGCGCGCACCCATCTGTTGGGGGTTCTGCGCGCAGACCAACCGGCGCCTCAC	1200
Db	1470	GTTGACGGGACACACCAAGTACAGGGGGAAGGGTACCTCCAGACCCAGAGCCTCGTG	1529
Oy	1201	AGCTTATTGACATGGGGCCCAAGGCAAAAAATCCAGCTCTGTTAACCCATATGGCACTGG	1260
Db	1530	TCCGGGCTCTCAAGAGGCCCATCTCAGAAAAATCCAACTCGTGAACACCAAGGAGCTGG	1589
Oy	1261	CACATCAACCGCACCGGCCCTGAACATGCAATAGACTCTTTGACACACCGGCTTATCGCGCT	1320
Db	1590	CACATCAACAGAGACCGGCTCTGTAATTTGCAATGACTCCCTCCAAACTGGGTTCAATTTCTGGC	1649
Oy	1321	CTGTTCTAACAACCAAGCTTCAACTCGTCAGAGATGTCCGAAGCGATGTGCGCTCGC	1380
Db	1650	CTGTTCTTACGACACAGAGTTCAACGCGCTCCGGGTGCGCAGAGCGCATGGCTAGCTGCGC	1709
Oy	1381	AGTATCGAGGCTTCCGGGTGGATATGGGGCCCTTGGCAATTTAGAGATATATGTCACAAT	1440
Db	1710	CCCATCGATGATTTGCTCAAGGGGTGGGTCCCATCACTCATGATATGCTCAGAGAGTC	1768
Oy	1441	CCAGAGATATAGAGACCCCTATTTGCTGGAGTACCACCAAGGCAAGTGTGGGTGCTCC	1500
Db	1769	-----GGAACAGAGGCCATATTTGCTGGCACTACGCGCTCAACCGTGGGGATGTGCT	1823
Oy	1501	GCGAAGACTGTGTGTGGCCCAAGTGTACTGTTTCAACCCCAAGCCGAGTGTATGTGGACG	1560
Db	1824	GCGTCCGAGGTGTGTGTGTCAGATGATTTGCTTCACTCCGAGCCCTGTTGTATGTGGGACG	1883
Oy	1561	ACCGACAGGCTTTGGAGCGGCCCACTTACACGTGGGGGGAGATATGAGACAGATGTCTTCTA	1620
Db	1884	ACCGATCTTTTGGGCTCTTACGTATGTGCTGGGGGAGAAATGAGACAGAGTGTCTGCTA	1943
Oy	1621	TTGAACAGACATCGACACACCGCTGGGGATCATATGTTCCGCTCAGCTGATGAATCTTCT	1680
Db	1944	CTTAGCAACACGCGGCCGCCCTTCAGAGCAATGTGTTTGGGTGCACGTGATGACACGACT	2003
Oy	1681	GAGTACACCAAGACTTGGCGGCGCACCACTGCGGTACTAGAGCTGACTTCAACGCGACG	1740
Db	2004	GAGTTTCAACCAAGACTGCGGGGGGCCCTCCGTGCAACATCGGGGGGGGTGCGCAACAAC	2061
Oy	1741	ACGACACTGTTGTGCCCCACGGACTGTTTATGAGAGCATCTGTATACCACTTACTCAA	1800
Db	2062	----CTTTGTCTGCCCCACGGAATGTTCGGAAGCACCCGAGGCCCACTTACCAAAAG	2117
Oy	1801	TGCGGCTCTGGGCGCTTGCGTCAAGCGCAAGTGCATGCTGACTACCCCTCAAGGCTTGG	1866
Db	2118	TGTGCTCGGGGCGCTGTGTGACACCCAGGTGCATGTGTGACTACCATACAGGCTCTGG	2177
Oy	1861	CATTACCCTGACAGTTTACTATVACATCTTCAAAATPAGATGATGTGGAGGGCTT	1920
Db	2178	CACCTACCCCTGACGTTTAACTTTTACCGTCTTTAAGTCAAGATGTATGTGGGGGCGTG	2237
Oy	1921	GAGCACAGGCTCAAGGCTGCAATTCATCTGTGGGGATTCGTTGCAACTTGGAGGAC	1980
Db	2238	GAGCACAGGCTCAATGTCTGCAATGTGAATTCAGAGAGAGCGCTGTGACTTTGAGGAGC	2297
Oy	1981	AGAGACAGAAATCAATGTCTCTCTTTGTTGCACTCACACAGGAATGGGCAATTTTACT	2048
Db	2298	AGGATATAGTCGAATCATAGCCCGGTGCTGTCTTACAAAGATGTGCAGATATCTGGCC	2355
Oy	2041	TGCTCTTACTGGAGCTGCGCGCTTGTGACTGTGTCTTTCACCTTCACCAAAATC	2100
Db	2358	TGTTCTCTTACCAACCTTACCGGCGCTGTTCACATGCTTGATTCATCTTACCGGAACATC	2417
Oy	2101	GTGAGCGTCAATTCATGATATGGCTATACCTGTGCCACAAATATACATGTCGAGATG	2160
Db	2418	GTGAGCGTCAATTCATTCAGGTATATAGGCTGGAGATGTGTCTCTTTGCAATCAAAATG	2477

QY	2161	GAGGAGGTAAPACTCTTATTCCTGCTCTTTCAGCGGACCGCAGGGTTTGCGGCTGTAATGG	2222
Db	2478	GAGATATATCCGTGTCTTTCTTTCTTCTTCTGCGGACCGCGCGCTCTGTGCTCTGTGG	2537
QY	2221	ATGCTCATCTCTGTGTGGCCAGGCGCGAGCAGCACTAGAGAGCTGTGCATCTTTCACGCT	2280
Db	2538	ATGATGCTGCTGATATGCCAGGCTGAGGCCACCTTATAGAAACCTGGTGGTCTCTCAATATGC	2597
QY	2281	GCGAGCCAGCTAGCTGCTCAATGGCTTCTTATTTTGTCACTTTTTCGTGGCTGTGG	2340
Db	2598	GCGCTGTGCGCGGAGCGCATGCGCTTCTCTCTCTCTGTGTCTTCTGCGCGCTGG	2657
QY	2341	TACATCAAGGGTGGGGTAGTCCCTTAGCTAACCTATCCCTCACTGGCTGTGGTCTT	2400
Db	2658	TACATCAAGGCGAGGCTGTGTCTGCGGCGCGCATATCTCTTATGGCGTATGGCGTTG	2717
QY	2401	AGCTTACTGTCTCTAGCATTTGCCCAACAGCTTATGCTTATGACGATCTGTGCATGGC	2460
Db	2718	CTCTGTGCTTGTGTGGCTTATACCAACCAAGCTTATCCATGGAACGAGATGGCTGCA	2777
QY	2461	CAGATAGAGCGGCTCTGCTGTATATATCACTCTCTTACTCTACCCCCGGTATAG	2520
Db	2778	TCGTGCGAGGCGCGGTTTTGTATAGTCTGTATCTCTTGACCTGTGTACCATATAG	2837
QY	2521	ACCTTCTCAGCGCGGTTTTGTGTGGTGTGGCATCTTGACCCCTGGGGGAAGCATG	2580
Db	2838	GTGTTCTCTCGTACCTCATATAGTGTGTACATATTTTATCACAAGCCAGACCGCAC	2897
QY	2581	GTCCAGAGATGGGACCACTATGCAAGTGCAGCGCGGTGCGGTATGATCATATAGGCC	2640
Db	2898	TTGGAATGTGGGTGCCCCCTCTCAATGTTTGGGGAAGCGCGCATGECATCATCTCTT	2957
QY	2641	GTGCCCATATTCTAACCAAGTGTGTGTGTGACATTAACAAGTGTCTTGGCGGTCTT	2700
Db	2958	ACATGCGCGGTCCATCCAGAGCTAATCTTTGACATCAACAACTCTGTCTGCGCATCTC	3017
QY	2701	GGGCTGCTTAACTCCCTAAAGGCTTTTGAACGGGTGGCGGTACTTGTCAAGGCTCAC	2760
Db	3018	GGTCCGCTCATGTGTCTCAAGCTGGGCATTAAGATGTGCGGTACTTTGTACGCGCTCAG	3077
QY	2761	GCTCTACTGAGATGTGCACCATGAGCAAGCATCTCGCGGAGGAGGTAACGTCCAGATG	2820
Db	3078	GGGCTCATCCGTGATGATGATATTTAGTGTGGAAAGTCCCTGTGAAGCCATATGTCCAAATG	3137
QY	2821	GCGCTACTAGCCCTTGGCAAGTGGACTGGCACTTATCATATGACACTTACCCCTATG	2880
Db	3138	GCTTTCATGAAAGTGTGCGCGGCTGACAGGTATCGTACGTATATGACACTTACTTCCACTG	3197
QY	2881	TCGGATTTGGTGTCTAGTGTGCTTCGGGAGCCTGCGCGTCCCGTTAGGCTTATCATTTTC	2940
Db	3198	CGGATATGGGCGCACGCGGCGCTTACAGAACCTTGGCGGTGGAGTAGACCCGTGTCTTC	3257
QY	2941	AGTCCGATGAGAGAAAGTCAATGTTGTGGGGACCGAGACAGTGTGTTGGGGACATT	3000
Db	3258	TCTGACATGAGACATTAACTCATACCTGTGGGGGCAACACCGCGGTGTGGGGACATC	3317
QY	3001	TTTACACGACTTCCCGTGTGCGCCGCACTTGTGTGGAGAGTCTCTTGTGCGCCAGCTGAT	3060
Db	3318	ATCTCGGCTTACCAAGTCTCCGCCGGAAGGGGGAAGAGATACCTTTAGACCGGCGGAT	3377
QY	3061	GGCTATACCTCCAAAGGGGTGAGTCTTTCGCCCCCATCACTGTTACGCCACACAGCA	3120
Db	3378	AGTTTGGAGAGCAGGGGTGGCGGCTCTTGGCGCTTATCAGCGGCTTATTTCCCAACAAACG	3437
QY	3121	CGTGGCCCTTTTGGGACCATATGTGGTGGATGACGGGGCGCGACAGAACGAAACAGGCT	3180
Db	3438	CGGGGCTTGTGGCTGTATATCATCTAGCTTCAAGGTGGGGACMAAGACCAAGTGCAT	3497
QY	3181	GGGGAAATTCAGTCTGTCCACAGTCACTCAGTCTTCTCGGAAACATCCATCTCGGG	3240
Db	3498	GGGAGAGTTCAAGTGTCTCCACCGCAACGAACTTTTCTGTGGGACCTGTGTCATATGGC	3557
QY	3241	GTTTTGTGAGCTGTACCATGAGTGGGCAACAAGACTTGGCCGGCTACCGGGGTCCG	3300

Db 3558 GTGTGTTGGAACCGTCTACATGATGTCGGGCTCGAAGACCTGGGCGGACCGAAGAGGTCCA 3617
Qy 3301 GTACGCAAGATGTAATCTCAAGTCTGAGAGGGGACTTAAAGAGTGGCCAGCCCCCTGGG 3360
Db 3518 ATACCCCAATGTAACCAATGTAACAGGACCTCGTGGCTGGCCGGCCCCCGGG 3677
Qy 3361 ACTAATCTTTGAGCGGTGACGATGTAAGCGGTGACCTGTACTGTCAACGCGGAC 3420
Db 3678 GCGGCTTCATGACACCGTGAACCTGGGACGCTCGACCTTAACTTGTGACAGGAT 3737
Qy 3421 GCTGATGTCATCCCGGCTCGAAGACGCGGGGCAAAAGGGAGGCTACTTCCCGAGA 3480
Db 3738 GCTGATGTCGTTCCGGTGGCGCGGGGGGAGACAGAGAGGGGAGACCTGTTTCCCGAG 3797
Qy 3481 CTTCTTTTCACCTTGAAGGGGCTCTCAAGAGGCCGGGTCTATGCCCCAGGGGCAAGCT 3540
Db 3798 CCGATCTCTAATCGAAGGGGCTCTCGGTGTAACACGTGTTTCCCTTGGGGGCAAGTT 3857
Qy 3541 GTGGAAGTCTTCCGGGAGGCTGTGCTCGGGGGCGGTAGTGCATAGTTGATC 3600
Db 3858 GTAGGACATCTTCCGGGCTGTGTGTGACCCGGGGGTTGGAAAGCGGTGACCTTACATA 3917
Qy 3601 CCGGTTGAGACACTGACATGTCACGCGGTCGCCACCTTAACTGACAAACAGACACCA 3660
Db 3918 CCGGTTGAGTCTAAGAACTACATGCGGTCTCGGCTTTCACAGACATCATCTCCCT 3977
Qy 3661 CTTGCTGTGCCACAGACTATCAGTGGGTACTTGCATGCCGACCTGACAGTGAAG 3720
Db 3978 CCGGCTTACCGCAAACTTCAAGTGGCAATTAAACAGCTCCACCTGGCAGCGGAG 4037
Qy 3721 AGCACCAAGTTCCTGCGGATATGCTCAAGGGATAAAGTCAAGTCTTAATCC 3780
Db 4038 AGCACCAAGTTCCTGCGGATATGCTCAAGGGATCAAGGCTGCTGCTTAATCCG 4097
Qy 3781 TCAGTGCTGCCACCTGAGGATTTGGGCGTACTTGTCTAAGGACATGACATCAATCC 3840
Db 4098 TCGGTTGCCGCAATTTGGGCTTTGAGCGTATATGTCAAAGGACATGACATGACCT 4157
Qy 3841 AACATTAGACTGAGTCAAGACTGTGACACCGGGCGCCCATCAGTACTCCACAT 3900
Db 4158 AACATTAGACTGAGGATTAAGGACCATCAACCGGGCGCCCATCAGTACTCCACAT 4217
Qy 3901 GGGAAATTTCTGCGGATGAGGGGCTGTGCGGGCGGCGCCCTACACATCATATATGAT 3960
Db 4218 TGAAGTTCTTGGCGGAGGTGTGCTCGGGGGGCGCTATGACATCATATATATGAT 4277
Qy 3961 GAATGCCATCCGTCGACTCAACCATCTTGGCATTCGGAACAGTCTTGTGATAGCA 4020
Db 4278 GAATGCCATCACTGACTGACTCAACATCTTGGGACATCGGACAGTCTGTGATAGCA 4337
Qy 4021 GAGACAGCTGGGCTCAGACTTAATCTGTGCTGCTACAGCTACGCCCCCTGGGTGATGACA 4080
Db 4338 GAGACAGCTGGAGGCGGCTGTGTGTGCTGCGCACCGCCATCGGATCATCAC 4397
Qy 4081 ACCCCCAACCCCAATAGAGAGTGGCCCTTGGGAGAGAGGCGGAGTCCCTTTAT 4140
Db 4398 GTGACACACCCCAATGAGAGTGGCCCTGTGCAACACTGAGAGATTTCTTTAT 4457
Qy 4141 GGGAGGCGATTCCTCTGTCTTACATCAAGAGAGAGATCTGATCTTCTGCAATTA 4200
Db 4458 GGGAGGCGATTCCTCAATGAGGCAATCAAGGGGGAAGGATCTCATCTTCTGCAATTC 4517
Qy 4201 AAGAAAAAGTGTACGAGCTCGGGCGGCGCTTGGGGTATGGCTTGAATCTCAGTGGCA 4260
Db 4518 AAGAAAAAGTGTACGAGCTCGGGCGGCAAGTGTCAAGGCTCGGACTCAATGTCTGTAGCG 4577
Qy 4261 TACTACAGAGGTTGAGCTCTCGTAAATACCACTAGGGAGAGCGTAGTGTGTGTGCC 4320
Db 4578 TATTTACGGGAGTCTCGATGTGTCTGTATACCACTAGCGAGAGCGTGTGTGTGTGGA 4637
Qy 4321 ACCGACGCTCTCATGACAGGATATCTGGGAGCTTTGACTCGGTGATGACTGCAACGTA 4380

Db 4638 ACAGAGCTCTAAATGACGGGTTTACCGCGGCACTTGTACTCAAGTATGACTGCAACACA 4697
Qy 4381 GCGTCACTCAAGTTGAGCTTCAAGTTGAGACCCCAATTCACATTAACACAGATT 4440
Db 4698 TGTGTACCCAGACAGTCAATTTCAAGTTGAGATCCCACTTCAACATTTAGACACAGCG 4757
Qy 4441 GTCCCTCAAGAGCTGTCTCAGCTAGCAGCGCGGGGATGCAAGGGTATAGGGAGACTG 4500
Db 4758 CTGCCCCAAGACCGGCTGTGCTGCGGACGCGGAGGATGAGTGTGACAGGGGAGAGT 4817
Qy 4501 GGCATTTATAGTATGTTTCACTGTGAGCAGGCTCAGAGAAATTTTGAACATGTATGTG 4560
Db 4818 GGCATTTACAGTTTGTGATCTCAGAGAGAAAGGCGCTCAGAGATGTTGATCTCTCGCTC 4877
Qy 4561 CTCTGTAGTGTACAGACGAGAGGCGGATGTGTATGACTCACACCATGAGAGACCAAC 4620
Db 4878 CTGTGTAGTGTATAGACGAGGCTGTGTGTATGATGTACGCGCGCTGAGACCTCG 4937
Qy 4621 GTGAGGCTCAGGGGATTTTCAACAGCGCGGTTTGTGTGTGCAAGACATCTTGAG 4680
Db 4938 GTTATGTTGGGCTTAACTAAATACAGAGGTTGCCGTCTGCAAGACACTTATAG 4997
Qy 4681 TTTTGGAGGCAATTTTCAACGCGCTTCAACATAGATGCCACTTCTTCCCAACA 4740
Db 4998 TTTCTGGAGAGCGTCTTCAAGGCTTCAACCATATGATGCCACTTCTTGTCCAGACC 5057
Qy 4741 AAGCAATCGGGGAAAAATTTGCAATCTTAACTTAACTTAACTTAACTTAACTTAACT 4800
Db 5058 AAACAGGCAAGAGAACACCTCCCTTACCTGTGTAGCATACCAAGCAGGTGTGCGCAG 5117
Qy 4801 GCCAAAGCCCCCGGCTGTGAGAGTATGAGAGGTTTGAATCTGACTCAAGCCC 4860
Db 5118 GCTCAGGCTCACTCTCATGTGTGAGACCAATGTGAGATGTCTATACGCTTAAGCCC 5177
Qy 4861 ACACTGTGAGCCCAACCTCTCTGTACCGGTTGGGCTCTGTAAACAGAGTACC 4920
Db 5178 ACACTGTATGGGCAAGCGCCCTGTGTGTACAGGCTAGAGCCGTTCAAAATAGAGTACT 5237
Qy 4921 CTCAACATCCGTCAGAAATATCATGCGCACTGTGATGAGAACCGACTTGTAGGTATG 4980
Db 5238 CTCAACACCCCAATACCAATATCATATGAGTATGATGTGAGTGTGAGTGTGCT 5297
Qy 4981 ACCAGACATGGGCTTGTGAGGGGAGTGTGAGGGGCGCTGCGCGGTATGCTGAGC 5040
Db 5298 ACTAGACCTGAGGCTGTATGAGGAGAGTCTTGGGCTTGGCGGCTTCTGCTGATG 5357
Qy 5041 ACCGGGTGTGTGATCATCTGCGCGCTTGCACATTAACAGAGGAGCGCTGTGCGCG 5100
Db 5358 ACAGGACGCGTGTATATTTGGGCAAGATCATCTTGTCCGGAGGCGCAGCTGTATCCC 5417
Qy 5101 GACAGAGAGTCTCTATGAGGCTTTTATGATGATGAGAGATGTGCTCTAGGGCGGCT 5160
Db 5418 GACAGAGAGTCTCTATACAGAGATGTGATGATGAGAGATGTGCTCTTCAACCTTCT 5477
Qy 5161 CTGATTGAAGAGGGGAGCGGATAGCGAGATGCTGAAGTCAAGATCAAGGCTTATG 5220
Db 5478 TACATGAGAGAGAAATGACAGTCTGCGGACAAATTCAAACAGAAAGGCGCTCGATTCG 5537
Qy 5221 CAGCAAGCTTTCAAACAGGCTCAAGACATTAACCACTGTGAGAGGCTTATGAGCCAG 5280
Db 5538 CAACAGGCTTCAGAGAGGAGGAGGCTGCTGCTCCGTTGTGAGTCAAGTGTGAGGCG 5597
Qy 5281 GTAGAAATATTTGGGCGCAACATGTGGAATCTTATAGGGCATCAATACCTGCA 5340
Db 5598 CTGAGAGTCTTGTGGGCAACATGTGGAATCTTATACGCGGATACAGTACTTGTGCA 5557
Qy 5341 GACATATCAACCTGCAAGGAAACCTGAGTATCTTCAATGAGGCTTCAAGTGTGAGC 5400
Db 5658 GGCCTATCACTGTGCTGAGAAACCCCGGATAGCATATGATGTGCTTTTACAGCTCT 5717
Qy 5401 CTGACAGTCTGCTGTCAACAGACCATATCTTCTCAACATTTTGGGGGCTGTGCTA 5460
Db 5718 ATCAGCAGCGCGCTCAACCAACCAATACCTCTCTTTTAACTTTGGGGGATGGGTG 5777

QY 5461 GCATCCCAATTTGACCAACCCGCGGGGCACTGGCTTTCGTTGTCAATGGCTTAGTGGGA 5520
Db 5778 GCTGCCCAACTCGCTCCCCCAAGGCGCTGTTCCGCTTTTCGTGGCGCGCATTTGGCGGT 5837
QY 5521 GCTGCCGTAGGCACTATAGGCTTAGTAAAGTGTCTAGTGAATCTCTGGCAGGGTATAGT 5580
Db 5838 GCGGCGGTGGACATAGTGTCTGGGAAGGTACTTGTGACATTTCTGGCGGGCTTAGGG 5897
QY 5581 GCGGGCAATTTTGGGGGGCTCTGTGCGCATTTCAAGATCATGTGTGGCGAAGACCTTCATG 5640
Db 5898 GCGGGGGGTGGCTGGCGCATCTGTGGCTTTAAAGTCATGAGCGCGAGATGCTCTCACT 5957
QY 5641 GAGGATGTCTCAACTTGTCTGCTGGAATTTCTCTCCGGGTGCTTGGTAGTGGAGTC 5700
Db 5958 GAGGATCTGTGTTAATTTACTCCCTGCGCATCTTTCTCTGGCGCGCTGGTGTGTGGGGTCT 6017
QY 5701 ATCTGCGGCGCATTTCTGCGCGCAACGTGGGACCGGGGGAAGCGCGCTTCATATGATG 5760
Db 6018 GTGTGCGGAGCAATATGCGGTGGGACGTGGGCGCGGGAGAGGGGGCTGTGCACTGGATG 6077
QY 5761 AATAGACTCATTTGCTTGTCTTCCAGAGAAATCACGTCGCGCCGACCACTAGTAGCG 5820
Db 6078 AACCGGCTGATAGCTTTCGCTTTCGCGGGGTAAACAAGTCTCCCGACGCACTATGTGCC 6137
QY 5821 GAGTCGATGCGGTGCGACGCTGTGACCCCACTATTGGCTCCCTTACCATTAACGAGCTG 5880
Db 6138 GAGAGCGACGCGCGGGCGGTGTATCTCAGATCCCTTCAGCTTACCATCATCTCAAGTTG 6197
QY 5881 CTCAGAAAGCTTCCACACTGGATTACTAGAGACTGCCCCCATCCATGCGGGGCTCGTG 5940
Db 6198 CTGAAGAGCTTCACTAGTGAATTAATAGGACTGTCCACGCTTGTTCGGGCTCGTG 6257
QY 5941 CTCGCGAGTGTGGGGACTGGGTTTGACACATCTTAACAGACTTTAAATAATTTGCTGACC 6000
Db 6258 CTTAAAGGATCTTTGGGACTGATATGACAGGTGTGGAGTGACTTCAAGACTTTGGCTCCAG 6317
QY 6001 TCCAAATTATTTCCCAAGATGCGCGGCTCCCTTTGTCTCTCTCAAAAAGGGGTACAG 6060
Db 6318 TCCAAAGCTCTGCGCGGTTACCGGGACTCCCTTTCTGTATGCAACGCGGGTACAG 6377
QY 6061 GCGGTGGGGCGGCACTGGCATCTAGACACACAGGTCTCTTGGCGGCCAATATCTCT 6120
Db 6378 GAGGTGTGGCGGGGATGGACATGTGCAACCACTGCGCCCATGTGGAGCAACATCAACC 6437
QY 6121 GGCATATCTCCGCTTGGGCTCCATGAGATCAAGGGGCTTAAGACTGTGATATATCTGG 6180
Db 6438 GGACATGTCAAAAAATGCTCTCATGAGATTTGTTGGGCCAAAAAATTGAGCAACAGTGG 6497
QY 6181 CAGGGGACCTTTCCTATCAATTTGTACACGAGGGCCAGTGCCTGCGGAAACCCGCGCA 6240
Db 6498 CATGGAACATTTCCCATCAACGATACACACGCGGCCCTTCGCAAGCTCCCGACGCGCG 6557
QY 6241 AACTTTAAGTGTCCATCTGAGAGGTGGCGGCTTCAGATGACGCGAGATGACCGACAC 6300
Db 6558 AACTATTTCCAGGCGCTGTGGGGGTGGCTGTGAGGATGACGTGAGATTAACCGGGTGG 6617
QY 6301 GGGTCATACCACTACATAACAGGACTCACCACTGATACTTGAAGTCCCTGCGCACTA 6360
Db 6618 GGGGATTTTCACTAATGTAAGGAGTGAACACTGACCAACGTGAATGCGCATGCGCAAGTT 6677
QY 6361 CCTCTCCCGAGTTCTTTTCTGGGTGACGGAAGTGCAGATCCATAGGTTTGCCTCCCA 6420
Db 6678 CCAGCCCTGAATTTTTCACGAGGTGGATGAGTACGTTGCAAGATATGCTCCAGTG 6737
QY 6421 CCGAAGCCGTTTTTCCGGGATGAGTCTCTGTTCTGCGTGGCTTAACTATTTGTGCTC 6480
Db 6738 TGCAGAACTCTCTACGAGAGAGGTGTATTTCCAGGTGTGGCTCAACAGATACCTGTCTC 6797
QY 6481 GGGTCCCAAGCTTCTTTGGCAACCTGACCCGACAGAGATATTGATGTCTCATGTATACA 6540
Db 6798 GGGTACAGCTCTCCATGTAGGCCGAAACCGAGTGTGCAAGTGTCTCATTTCCATGCTCAC 6857

QY 6541 GATTCATCTCATATCAACGCGGAGACTGACGCGCGGCTTTAGCGCGGGGTTACCCCA 6600
Db 6858 GACCCCTCTCATATTAACAGACAGACGCGCAAGCGTAAAGCTGGCCAGGGGGTCTCCCGCC 6917
QY 6601 TCCGAGGCAAGCTCCCTCGCGAGGCGAGCTATGGGACATCCCTGTGAGACCCCTGACCC 6660
Db 6918 TCCTTGGCGAGCTCTTCACTAGCCAGTGTCTGCGCTTCTTTTGAAGGACATGTACTT 6977
QY 6661 ACCCAGGCAAGACCTATGATGTGACATGATGATGCTAACT-----GTTC 6708
Db 6978 ACCCATCATGACTCCCGGACGCTGACCTCATGAGGCCAACCTCTGTGGCGGACGAG 7037
QY 6709 ATGGGGGGGAGATGTGACTGATAGACTCTGGGTCCAAAGTGTCTTTGACTCTCTC 6768
Db 7038 ATGGGCGGGAACATCAACCGGTGTGAGTGAAGAAATAGGTGATATCTTGAATCTTTTC 7097
QY 6769 GACCCAAATGATGGAAGAGAGCGACTTGAAGCTTGTGATACATCAGATATCATGCTC 6828
Db 7098 GATTCGATTTGCGCGGTGGAGATGAGAGGAATATCCGTCCCGCGAGATCTTGCA 7157
QY 6829 CCCAAGAGAGTTCCCAACAGCTTTACCGGCTGGGCAACGCGCTGATTTACACCCAG 6888
Db 7158 AATCCAGAGAAATTCGCCAGGCTTGGCCATATGGGCAACGCCGATTTACAACTTCCA 7217
QY 6889 CTGTGGAATCTGTGAAAAGCCAGATTACCAACCGGCACTGTTGCGGGCTGTGCTCTC 6948
Db 7218 CTGCTAGAGTCTGTGAAGACCGGACTAGTCCCGCGGTGTACCGGGTGTCCCTTTTG 7277
QY 6949 CCTCTCTTAGAAAACCCCGACGCTCCCGCAAGAGGCGCGGACAGTGGGCTTAAGT 7008
Db 7278 CCACTTACCAAGGCCCCCAATACCACTCCACGAGGAAGAGCGGTTTCTTGACA 7337
QY 7009 GAGACTCCATAGAGATGCTCTTCAACAGCTGGCCATTAAGTCTTTTGGCAGCCCCC 7068
Db 7338 GAGTCCACGCTGTCTTCTGCTTGGCGAGCTGTACTTAAGACTTTTGGCAGCTCGGG 7397
QY 7069 CCAAGCGGCGATTCAGGCTTTTCAACGGGGGCGGGGCTGCGGATTCGCGACTGAGAG 7128
Db 7398 TCGTCGCGCGTTGACAGCGG---CACGGGCACTGGGCGCTTCCGCA-----TCAGGCC 7445
QY 7129 CCTCTGATGATGTGGCCCTTTTGGAGACAGGTTCCATCTCTTCAATGCCCCCTCGAG 7188
Db 7446 TCCGACGACGCGCAAAAGGATCGACGTTGAATGCTATCTCTCATGCCCCCTCTCGAG 7505
QY 7189 GGGGAGCTTGGAGATCCAGACTGAGCGCTGAGCAGATGAGAGCCCAACCCCCCCAG 7248
Db 7506 GAGAGGCGAGGGGACCCCGACT----- 7528
QY 7249 GGGGGGGTGGCAAGCTCCCGCTGAGCTGGGGTCTTGGTCTATCTTGTCTCCAGAGAGAC 7308
Db 7529 -----CAGCGACGGGTCTTGTCTTACCGTGAACGGGGAAGCTGGT 7568
QY 7309 GACTCGGTGTGTCGCTCCATGTCATATCTCTTGAACGGGGCTCTTAATTACTCTTGT 7368
Db 7569 GAGGACGTGTCTGTGCTCTCAATGTCTTATATGACAGGTGCTTTGATCAGCGCATGC 7628
QY 7369 AGTCCGAGAGAGGAGATTACCGATTAACCCCTTGAAGCACTCCCTGTTGCGATATCAC 7428
Db 7629 GCTGCGGAGGAGAGCAAGTTGCCATCAATCCGTTGAGCAACTTTTGTCTGCCTCACAC 7688
QY 7429 AACAGGTGTACTGTACACACAAAGAGGCTCTACTAAGGCTTAAAAAGTAACTTTT 7488
Db 7689 AGTATGTTCTATCCCAACATCTCGACGGCAAGCTGTGGGAGAGAAAGTCACTTTT 7748
QY 7489 GATAGATGCAAGTGTGCACTCTTACTAGCACTCAGTCTTAAGGACATTAGCTAGCG 7548
Db 7749 GACGAGCTGCAAGTCTTGGAGACACTTACGGGACGTGCTCAAGAGATGAAGGGAAG 7808
QY 7549 GCCTCCAAAGTCAACGCAAGCTCTCAACATGAGAGAGGCTTTGCAATTAACTCCACCC 7608
Db 7809 GCGTCCCAAGTTAAAGGTAGGCTTTATCTATAGAGAGGCTGTGCAATGACGCCCA 7868
QY 7609 CATTTGCAAGATCTAATATAGGTTTGGGGCTTAAGAGGTCCGCAAGCTTGTCCGGAGG 7668

CC detecting method of the invention. (Updated on 27-OCT-2003 to standardise
CC OS field)

Sequence 9413 BP; 1887 A; 2815 C; 2694 G; 2017 T; 0 U; 0 Other;

Query Match 45.5%; Score 4140.2; DB 8; Length 9413;
Best Local Similarity 66.6%; Pred. No. 0;
Matches 6065; Conservative 0; Mismatches 2953; Indels 93; Gaps 6;

```
QY 1 ATGAGCAAAATCTTAAACCTCAAGAAAACCAAAAAGAAACCAACCAACCGTCCGCCA 60
DB 330 ATGAGCAAAATCTTAAACCTCAAGAAAACCAAAAAGAAACCAACCAACCGTCCGCCA 389
QY 61 GACGTTAAGTTCCGGGCGCGCGCCAGATCGTTGGCCGAGATATCTTTGGCCGCGAG 120
DB 390 GACGTTAAGTTCCGGGCGCGCGCCAGATCGTTGGTGAATTTACCTGTTGCCGCGAG 449
QY 121 GGGCCCAAGTTGGGTGGCGCGCAAGAAAGACTTCGAGCGGTCCAGCCAGTGA 180
DB 450 GGGCCCAAGTTGGGTGGCGCGCAAGAAAGACTTCGAGCGGTCCAGAACTCGTGA 509
QY 181 AGGCGCAGCCATCCCTAAGATCGGGCTCCACTGGCAAAATCCTGGGAAAAACAGA 240
DB 510 AGGCGCAGCCATCCCTAAGATCGGGCTCCACTGGCAAGACTCGGCTCAGCCGCG 569
QY 241 TACCCCTGCCCCCTAATACGGGAATGAGGGAATCGGCTGGCGAGGATGATCTTCC 300
DB 570 TACCCCTGCCCCCTAATACGGGAATGAGGGAATGAGGATGATGATCTTCTGTCCC 629
QY 301 CGAGTTCCGCTCCCTCTTGGGGCCCCCAATGACCCCGGCAATAGTCCGCAAGTGG 360
DB 630 CGTGGCTCCGCTCTTGGGGCCCCCAATGACCCCGGCAATAGTGGCTGAATTTGG 689
QY 361 AAGGTCATCGATACCTTAACGTTGCGGCTTGGCGACCTCATGGGTTATCCCTGTG 420
DB 690 AAGGTCATCGATACCTTAACGTTGCGGCTTGGCGACCTCATGGGTTATCCCTGTG 749
QY 421 GGGCCCCGCTCGGCGCGCTCGCGAGACTCTCGCGCATGAGCTGAGAGTCTTGAAGAC 480
DB 750 GGGCCCCGCTCGGCGCGCTCGCGAGACTCTCGCGCATGAGTGTCCGGGTTCTGAAGAC 809
QY 481 GGGGTTAATTTTTCGAAACAGGGAATTTACCGGTTGCTCTTTTCTATCTTCTGCTGC 540
DB 810 GGGGTTAATTTTTCGAAACAGGGAATTTACCGGTTGCTCTTTTCTATCTTCTGCTGC 869
QY 541 CTGCTGCTCGATACCAACCCCGGCTCGGCTGCGAAGTGAAGAAATCATGATACCGGC 600
DB 870 TTGCTGCTGTTTGGACCATCCAGCTTCGCTTACGAGTGGCAACGATCGTCCGGA 929
QY 601 TACATGCTGAATGACATGACCAATGACAGATTAATGAGTGGAGCTTCAGGCTGCTG 660
DB 930 TACATGCTGAATGACATGACCAATGACAGATTAATGAGTGGAGCTTCAGGCTGCTG 989
QY 661 CTCCACGCTCCCGGGGTGCTCCGTTGCGAAGAAATGGGGAATGATCTCACTGCTGATA 720
DB 990 ATGACACACCCCGGGGTGCTCCGTTGCGAAGAAATGATCTCCCGTTGCTGATA 1049
QY 721 CGGCTCACGGAATGAGCGGTGAGCGGCGCGCGCTCACGAGAGGCTTGGGAGC 780
DB 1050 GCGCTCACGCTCCAGCGTCCGCGCAAGAAACAGAGCATCCCAACAGACAAATACGAGC 1109
QY 781 CACATGACATGTTGTGATGTCGCGCAAGCTTGTCTGTGCTCTTACGTTAGGGAGCTC 840
DB 1110 CACGTCGATTTGCTGTTGGGGCGGCTGCTCTGTGCTCCGCTATGTAAGTTGGGATTC 1169
QY 841 TCGGATGAGGATGATGCTGAGCCCAATTTTCAATTTGCTGCGCGCAACCACTGTTT 900
DB 1170 TCGGATGAGGATGATGCTGAGCTCCAGCTGTTCACTTCTCACTCCGCGATGAGAGC 1229
QY 901 GTCCAAAGTGAATTTGCTCATCTACCTGATACATCATGAGCAACGAGTGGCATGG 960
DB 1230 GTCCAAAGTGAATTTGCTCATCTACCTGATACATCATGAGTGAAGTCAACGAGCTGG 1289
```

```
QY 961 GACATGATGATGACTGTGCGCCACCGGCTACATGATCTTGGCGTACGATGCGTGTG 1020
DB 1290 GATATGATGATGACTGTGCGCCACCTAACAACCGGCTTGTGTATCCAGCTACTCCGATC 1349
QY 1021 CCCGAGTCAATTAATGACATCAATTAGCGGGCTCATTTGGGGGCTCATGTTGGCTTGGC 1080
DB 1350 CCACAGCGGTGTGACATGATGAGGCGGCGCCACTGAGGTTGCTTACGCGGCTTGTG 1409
QY 1081 TACTTCTATGACAGGAGCGTGGGGAAGTGTGTTGATCTCTTGTGAGCGCGG 1140
DB 1410 TACTTCTATGATGAGGAGCTGGGGAAGTGTGTTGATCTCTTGTGAGCGCGG 1469
QY 1141 GTGAGCGCGCACCCCACTGTTGGGGTTCGCGCGAGACCAACGAGCGCTCAC 1200
DB 1470 GTGAGCGCGCACCCCACTGAGGAGGAGGTTAGCTTCAGGACCCAGAGCTTCGTG 1529
QY 1201 AGCTTATTTGACATGGGCGCCAGGAGAAATTCAGCTGTGTTAACCAATGGAGCTGG 1260
DB 1530 TCTGGCTCTCAAAAGGCCCATCTTGAAATTCAACTGTGAACCAACGAGCTGG 1589
QY 1261 CACATCAACCGACCGCCCTGATGACATGACTCTTGGCACACCGGCTTTATCGGCT 1320
DB 1590 CACATCAACCGACCGCTGATGATGACATGACTCTTGGCACACCGGCTTTATCGGCT 1649
QY 1321 CTGTTCTTACCAACCAAGCTTCAACTGCTCAGAGATGCTCCGAAACGATGCTCGCC 1380
DB 1650 CTGTTCTTACCAACCAAGCTTCAAACGCTCGGAGTGCAGAGCGATGCTAGCTGCGC 1709
QY 1381 AGTATGAGAGCCCTCCGGGTGGGATGGGGGCGCTTGCAATATGAGGATTAATGACCA 1440
DB 1710 CCGATGATGAGTGTGCTCAGGAGGTGGGTCCATCATCATGATATGCTTGAAGCTC- 1768
QY 1441 CCAGAGATATGAGACCTATTGCTGGACCTACCAACCAAGCAGTGGGCTGTGCTC 1500
DB 1769 -----GACCAAGAGGCAATATTGCTGGACCTACCAAGCAGTGGGCTGTGCTC 1823
QY 1501 GCGAAGCTGTGTGGCCCAAGTGTACTGTTTCAACCCCAAGCCAGTGTAGTGGAGC 1560
DB 1824 GCGTCCCAAGTGTGTGGTCAAGTGTATGCTTCACTCCAGGCTGTTGTAGTGGAGC 1883
QY 1561 ACCGACAGCTTGGAGCGGCGCACTTACAGTGGGGGGAATGAGACAGATGCTTCTTA 1620
DB 1884 ACCGATGCTTGGAGCGGCTCTTACAGTGGGGGGAATGAGACAGATGCTTCTTA 1943
QY 1621 TTGAACGACTGACCAACCGCTGAGGATGATGTTGCGCTGACGATGATGATCTTCT 1680
DB 1944 CTTAGCAACAGCGGCGGCTCAAGGCACTGTTTGGGTGACGATGATGATGATGATC 2003
QY 1681 GGTACACCAAGACTTGGCGGACCAACCTGCGATCTAGAGCTGATTTCAACGCGAC 1740
DB 2004 GGGTTACCAAGAGCTGGGGGCGCTCCGCTGCAACATGAGGGGCTGGCAACA-- 2061
QY 1741 ACGGACGTGTTGGCCCAACGAGCTGTTTGAAGACATCTGATACCACTTACCTCA 1800
DB 2062 ----CCTGATCTGCGCCACAGATGCTTCCGAAACCCGAGGCACTTACCAAG 2117
QY 1801 TGGGCTTGGGCTGCTGCTGCTCAAGGAGTGTGCTGATGATGATGATGATGATGAT 1860
DB 2118 TGGGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2177
QY 1861 CATTAACCTGCAACGATTAATTAATCACTTCAAAATTAAGATGATGATGATGATGAT 1920
DB 1921 GAGACAGGCTCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
QY 2238 GAGACAGGCTCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2297
DB 1981 AGAGACAGAGTCAAGTGTCTCTTTGTTGACCTCAACCAAGGATGAGGCACTTTA 2040
QY 2298 AGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2257
DB 2041 TGCTTTACTGGAACCTGCGGCTTGTGATGATGATGATGATGATGATGATGATGATGAT 2100
```


Db 2258 TGTTCCTTCAACACCTTACCGGCGCTGTCTCACTGGCTGATCATCTTACCGGAACATC 2417
Qy 2101 GTGACGCTACATTCATGTATGAGCCTTATCACTGCGCTTACAAATATCATGTCGATGG 2160
Db 2418 GTGACGCTGCAATCTGTACGTATAGGCTGCGAGTGTCTCTTGTGCAATCAAAATGG 2477
Qy 2161 GAGTGGGTAATACCTTATCTGCTCTTACGGGACGCGAGGTTTGGCGCTGTATGG 2220
Db 2478 GAGTATATCTGTTGCTTCTTCTTCTTCTGCGGACGCGCGCTGTGCTGTGCTTGTGG 2537
Qy 2221 ATGCTCATCTTGTGGGACAGCGGACAGACACTAGAGAAAGTGTCTATCTGTGACAGCT 2280
Db 2538 ATGATGCTGTGATAGCCAGGCTGAGGCACTTAGAGAACTGTGTGTCTCTCAATGCG 2597
Qy 2281 GCGAGCGAGCTACTGCAATGAGCTCTTATATTTTGTGATCTTCTTGTGAGCTGTGG 2340
Db 2598 GCGTCTGTGGCGGAGGCGATGGCGCTTCTCTCTCTGCTGTCTTCTGCGCGCTGG 2657
Qy 2341 TACATCAAGGGGTGGGAGTAGTCCCTTAGCTACTATTCCTCACGTGGCGTGGTCTTT 2400
Db 2658 TACATCAAGGAGGAGGCTGTGTCTGTGGGCGGATATGCTCTATGGGATAGCGCTTG 2717
Qy 2401 AGCTTACTGTCTTACGATTTGCCCAACAGGCTTATGCTTATGACGCACTGTGCAATGGC 2460
Db 2718 CTCCTGCTCTTGTGGCTTACCAACAGACTTATGCAAGACGAGATGGCTGCA 2777
Qy 2461 CAGATAGAGCGGCTGTGCTGTATATGATCACTCTTTTACTCTACCCCGGATTAAG 2520
Db 2778 TCGTGGGAGGCGGCTTGTGTGTAGGTCTGTACTTGTACCTTGTACCATTAATTAAG 2837
Qy 2521 ACCCTTCAAGCGGTTTGTGTGGTGTGTGATCTCTGACCGCTGGGGAAGCTATG 2580
Db 2838 GTGTTCCTCGCTAGGCTCAATATGTGTGTACAAATTTTATCACAGAGCGGAGCGAC 2897
Qy 2581 GTCCAGAGTGGGACCAACCTATGCAAGTGGCGGCTGTGATGCAATATGGGCC 2640
Db 2898 TTGCAATGTGGGTCCCCCTCTCATGTTCGGGGAAGCGCGATGCAATCTCTCTT 2957
Qy 2641 GTGCCAATTTTCAACCAAGTGTGTGTGTAATTAACCAAGGCTCTTGGCGGTGTT 2700
Db 2958 ACATGCCGCGCTCATCCAGAGCTAATCTTGTGACATCAACAACTCTGTGCAATCTC 3017
Qy 2701 GGGCTGTCTTACCTCCAAAGGTGTGACGGCGCTGCGCTTCTGTCAAGGCTCAC 2760
Db 3018 GGTCCGCTCATGTGTCTCAAGCTGTGCAATTAAGAGTCCGTATTTGAACGGCTCAG 3077
Qy 2761 GCTTACTGAGATGTGCAACATGGCAAGCATCTCGCGGGGACAGGTATGTCAGATG 2820
Db 3078 GGGCTCATCGGTGATGCAATTTAGTGTGGAAGGTGCTGTGAGGCACTATATGCCAATG 3137
Qy 2821 GCGCTACTAGCCCTTGGCAAGTGTGACTGGCACTTAACTATATGCACTTCAACCCCTATG 2880
Db 3138 GCGCTTCAATGAAGCTGGCGCGCTGACAGTATGTAATGACATCTTACTCACTG 3197
Qy 2881 TCGGATTTGGGCTGTACTGTGCTGCGGGAACCTGCGGCTGCGCTTGAAGCCATATCTTC 2940
Db 3198 CGGATTTGGGCGCCACGCGGCTGACGAGACCTTGGGTGCGAGTAGAGCCGCTGTCTTC 3257
Qy 2941 AGTCCGATGGAGAAAGATCATTTGTCTGGGAGCGGAGACAGTGTCTTGTGGGAGATTT 3000
Db 3258 TGTGACATGAGAGCTAACTCATCACTGTGGGGGCAACACCGGGGTGTGGGGAATC 3317
Qy 3001 TTACACGGAATTTCCGTGTCCGCCGACTTGTGTGGAGGTCTCTTGTGCCAGCTGAT 3060
Db 3318 ATCTCGGCTTACCAAGTCTCCGCCGAAGGGGAAAGATACCTTCTTAGGACCGGCGAT 3377
Qy 3061 GGGCTATCTCCAAAGGGGTGAGTCTTCTGCGCCCACTGCTTGAAGCCAGACACA 3120
Db 3378 AGTTTGGAGAGCAGGGGTGGCGGCTCTTGGCTTATACAGGCTTATTTCCCAACAAAG 3437
Qy 3121 CGTGGCTTTTGGGACCATATGTGTGAGTATGACGCGGCGCGACAAAGACAGAGCT 3180

Db 3438 CGGGGCGTGTGGCTGTATCATCATAGCTCAAGGTGGGACAAAGACCAAGTGGAT 3497
Qy 3181 GGGAAATTAAGGTCTCTGTCCACAGTCACTCATGTCCTTCTGTGGAATTCATCTCGGG 3240
Db 3498 GGGAGGTTCAAGGTCTCTCCACCGCAACGAACTTTCCTGTGGACCTGTGTCAATGGC 3557
Qy 3241 GTTTTGTGAGCTGTCTCAATGAGCTGGGCAACAAAGCTGTGGCGGCTGACGGGGTCCG 3300
Db 3558 GTGTGTTGACCGTCTTACATGATGTGCGGCTCGAAGACCTTGGCGGCGCAAGGATCA 3617
Qy 3301 GTCAAGAGATGTACTCCAGTGTGAGGGGACTTATAGGTTGGCCAGCCCTCTGGG 3360
Db 3618 ATCAACCAAAATGTACCAATGTATGACAGGACTCGTGGCTGGCGGCGCCCGCGGG 3677
Qy 3361 ACTAAATCTTTGAGACCGTGTGACGTGTGAGACGGTGTCACTGTATCTGTGTACGGGAC 3420
Db 3678 GCGGCTTCATGACACGTTGCACTGTGGGACGTGGAACCTTATCTTGTGTACAGAGCAT 3737
Qy 3421 GCTGATGTCAATCCCGGCTCGAAGACGCGGGGACAAACGGGAGGCGCTTACTCTCCGAGA 3480
Db 3738 GCTGATGTCTTCTCGGTGCGCGCGGCGGAGACAGAGGGGAGCTGTCTTCCCGAGG 3797
Qy 3481 CCTCTTTCACCTTGAAGGGGTCTCTAGAGAGCCGGTGTATGCCCCAGGGGACAGCT 3540
Db 3798 CCATCTCTTACCTGAAGGGGTCTCTGGGTGTGACACTGTGCTTGGCTTGGGGACAGTT 3857
Qy 3541 GTCGAGTCTTTCGGGACGCTGTGTCTCTTCTGGGGCGTGGCTAATGTCATATGATTTATC 3600
Db 3858 GTAGGCAATCTTCCGGGCTGTGTGTGCAACCGGGGGGTTCGAAGGGGTGTGACTTATA 3917
Qy 3601 CCGGTTGAGACATCTGCAATGTGACAGGATCCCGGACCTTTAGTGAACAGACAGCA 3660
Db 3918 CCGTTATGTCTATGAGAACTACATATGAGGTCTTCGCTTTCACAGACAACTATCCCT 3977
Qy 3661 CTGTGTGTGCCAGACCTATCAAGTGTGGGTATCTTGATGCCCCAGCTGTGAGTGAAG 3720
Db 3978 CCGGCGGTACCGGAAATCTCAAGTGGCACTTATACGCTCCACTGGAGCGGAG 4037
Qy 3721 AGACCAAAATTTCTGTGTGATATGCTGTCAAGGGATTAAGTGTATGTATATCC 3780
Db 4038 AGCACCAAAATGTGGGCTGTGTATGCAAGGCCAAGGGGTACAAAGGTGTGTCTTAAACCG 4097
Qy 3781 TCAGTGTGTGCAACCCGTGGGTTTGGGGGTGTACTGTGTAAAGGACATGTGATCAATCCC 3840
Db 4098 TCGTGTGCGCACATTTGGGCTTGTGAGCTTATATGTGCAAGGACATATGATGTGACT 4157
Qy 3841 AACATTAGACTGAGTCAAGTCAAGTGTGACAGCGGGCGCCATCACTACTCCACATAT 3900
Db 4158 AACTCAGAACTGGGGTAAAGACATCAACACGCGGCGCCCATCACTCACTCAT 4217
Qy 3901 GGCAAATTCCTCGCGATGGGGGCTGTGCGGGCGGCGCTTACGACATCATATGTGAT 3960
Db 4218 TGCAAATTTCTTCCGCGGTGTGATGTCTCGGGGGCGCTTATGACATCAATATATGTGAT 4277
Qy 3961 GAATGCCATGCGGTGACCTTACCAACCATCTTGGGATGTGGAAACAGTCTTGTATCAAGCA 4020
Db 4278 GAATGCCATCAATGATCTGTGACATCAATCTTGGGATGTGGCAACAGTCTGTGATCAAGCA 4337
Qy 4021 GAGACAGCTGGGGTCAAGCTTAACTGTGTGGCTTACAGCTTACGCCCCCTGTGGGTCAAGTGA 4080
Db 4338 GAGACGCTGTGAGCGGGGCTGTGTGTGTGTGCGCACCGGCGCTCGGGATGATCACC 4397
Qy 4081 ACCCGGCAACCAATAGAGAGTGTGCGGCTTGGGAGGAGGCGGATCTCTTAT 4140
Db 4398 GTGCAACACCCCAACATGTGAGGAAGTGGCCCTGTCCAAACCTGAGAGATTTCCCTCTAT 4457
Qy 4141 GGGAGGGCAATCCCGTCTTATCATCAAGGGAGAGACATCGATCTTCTGCAATTA 4200
Db 4458 GGCAAAGCATCTCCATTTAGGCGCATCAAGGGGGAAGGATCTCATCTTCTGTGCAATTC 4517
Qy 4201 AAGAAAAATGTAGAGAGTGTGCGGCGGCTTTCGGGGTATGGGCTTGAATCTAGTGA 4260
Db 4518 AAGAAAGATGTAGCGAGTGTGCGGCAAGCTGACAGGCTCTGGACTCAATGTGTAGCG 4577

4261 TACTACAGAGGGTTGACGCTCCGTATATCACTGAGGAGAGCTAGTGTGTCGCC 4320
4578 TATTACGGGGCTTCGATGTGTCCGTATACCGACTAGCGGAGACGTGTTGTGTGGCA 4637
4321 ACCGACGCCCTCATGACAGGGTATATCTGGGGACTTTGACTCTCGTATGACTGCAACGTA 4380
4638 ACACACGCTCTTAATGACGGGTTTATACCGGCGACTTTGACTCAGTATGCACTGCAACACA 4697
4381 GCGGTACCTCAAGTTGTAGACTTGAAGTATGACCCGACATTCACCATTAACCAACAGATT 4440
4698 TGTGTACCCAGACAGTGTGATTTGACCTTGGATCCACCTTCACCATTTGAGACACACACG 4757
4441 GTCCCTCAAGACGCTGTCTCAAGTACGCGCCGAGGTGCGACGAGGTAGGGGAGACGCTG 4500
4758 CTGCCCCAAGACGCGGTGTCCGTGTGCGACGCGGAGGTAGGACTGGAGGCGGACGAGAT 4817
4501 GGCATTTATAGTATGTTTCCACTGTGTGAGCGAGGCTCAGGAATGTTTGAAGTGTAGTG 4560
4818 GGCATCTACAGGTTGTGACTTCAGAGAAAGCGCCCTCAGGCAATGTTGACTCTCGGTG 4877
4561 CTCTGTAGTGTCTCAAGCAGAGGGGCGCATGTATGAGCTCAGACCATGGAGACGACC 4620
4878 CTGTGTAGTGTCTATGACGAGGCTGCGCTTGTATGAGCTCAGCGCCGCTGAGACCTTCG 4937
4621 GTCAAGCTCAGGGCGTATTTGACACGCGCCGTTTGTCTGTGTGCAAGACCATCTTGAG 4680
4938 GTTAGGTTGGGGCTTACTTAATACACAGGGTTGGCCGTCTGCGACGAGACCATCTAGAG 4997
4681 TTTTGGAGAGCAATTTTCAACGGGCTCAGACATATAGTCCCACTTCTTCCCAACA 4740
4998 TTCTGGGAGAGCGTCTTCAAGGCTCACCCACATAGATGCGCACTTGTGTCCCAAGCC 5057
4741 AAGCAATCGGGGAGAAATTTGCACTACTTAACGCTTACAGGCTCAGGCTACAGTGTGCTAGG 4800
5058 AAACAGGAGAGACACCTCCCTACTCGGTAGCATACCAAGCCACAGTGTGCGCCAGG 5117
4801 GCCAAGCCCCCCCCGCTCTGTGGACGTCATGTGGAAAGTGTGACTCGACTCAAGCCC 4860
5118 GCTCAGGCTCCACCTCCATCGTGGGACCAAAATGTGGAAAGTGTCTCATACGAGCTAAAGCCC 5177
4861 ACATCTGTGGGCCCCCAACCTCTCTGTACCGCTTGGGCTGTGTTTCCAAAGAGTCAAC 4920
5178 ACATCTGTGGGCCCCCAACCGCTCTGTGTACAGGCTAGGAGCGCTTGTCAAAAGAGTCACT 5237
4921 CTCACATCCCGTGAAGAAATATACGCGCACTGTGCATGCAAGCGACCTTGAAGTATG 4980
5238 CTCACACACCCCATTAACCAAAATATACATGCGATGATGCGGCTGACCTGGAAGGTGTC 5297
4981 ACCAGCATGGGTCTTGGCAGGGGAGCTTGGCGGCGCTGCGCGCTATTTGCTGCGG 5040
5298 ACTAGCATCTGGGTCTAGTAGGGGAGGATCTTGGCGGCTGTGGCGGCTGATCTGCTGAGG 5357
5041 ACCGGGTGTGTTTGCATCATGCGCGCTTGCACATTAACGAGCGAGCGGTGTGGCGCG 5100
5358 ACAGGACCGTGTGATGATGTGGGACGAGATCATCTTGTCCGGGAGCGCACGCTGTTATTTCC 5417
5101 GACAAAGAGGTCTCTATAGGCTTTTATGATGATGAGAGATGAGCTCTTAAAGGCGGCT 5160
5418 GACAGGAGAGTCTCTTACAGGAGTTCGATGAGTGAAGAGTGTGCTTCCACACTCTCT 5477
5161 CTGATTTGAAGAGGGCAGCGATAGCGAGATGCTGAAGTCCAGATCCAAAGCTTATTTG 5220
5478 TACATCGAGCAAGAAATGAGTCTGCGCGAGCAATTCAAACAGAAAGGCGCTCGGATTTGCTG 5537
5221 CAGCAAGCTTCCAAACAAAGCTTAAGACATACAAACCACTGTGACAGGCTTACGCGCAAG 5280
5538 CAACAGGACCAACCAAGCAAGGAGCTGTGCTCCGTTGTGAGTTCCAAGTGTGGAGACC 5597
5281 GTAGAACAATTTCTGGGCGCAACACATGTGGAACTTATTAAGGCGATCCAAATACCTGCA 5340
5598 CTTGAGGTCTTCTGGGCGAAACACATGTGGAACTTATTAAGGCGATTCAGATTTGTGCA 5657

5341 GGAATATCAACACTGCGAGGAAACCTTCGAGTAGCTTCATGATGCGTTACGTGCGCC 5400
5658 GGCCTATCCACTGTGCTGGAAACCCGCGATAGCATATTAATGAGCTTTTACAGGCTCT 5717
5401 CTACCAAGTCCGTGTCAACAAAGCACATATCTTCTTCAACATTTTGGGGGGCTGGCTTA 5460
5718 ATCACAGGCCGCTCAACCAACCAATATCCCTCTGTTTAAACATTTGGGGGAGATGGAGT 5777
5461 GCATCCCAATTTGACACACCCGCGGGGCGCACTGAGCTTGTGTGAGTGTAGTGGAGA 5520
5778 GCTGCCCACTGCTCTCCCGACGCGCTGCTTGGCTTGTGGGCGCGCGCATTTGCGGCT 5837
5521 GCTGCGGTAGGACATATAGGCTTAAAGTGTAGTGAACATCTTGCAGAGGATAGT 5580
5838 GCGGCGGTGTGGCAGCATATGAGTCTCGGGAAAGTACTTGTGGAATTTGCGCGGAGTATGG 5897
5581 GCGGGCATTTGCGGGGCTCTCGCGCATTCAGATATCATGTGCGAGAAAGCCCTCCAGT 5640
5898 GCGGGGATGCGTGGCGCACCTGTGCTTTAAGTATGACGCGGAGATGCTTCACT 5957
5641 GAGATGTGCTCACTTGTGCTGCGGAATTTGTCTCCGAGTGTGCTTGTAGTGGAGTTC 5700
5958 GAGGATGTGTTATTTACTCTCTGCGCATCTTCTCTGCGCGCTGTGTTGTGGGGTTC 6017
5701 ATCTGCGCGGCAATTTGCGCGGACAGTGGGACGCGGAGAAAGCGCTCCATGAGATG 5760
6018 GTGTGCGAGCAATTAATGCGTGGGACAGTGGGCGCGGAGAGGGGCTGTGCAAGTGGATG 6077
5761 AATAGACTATTTGCTTCTTCCAGAGAAATACGTGGGCGCCCAACCACTACGTAAGAG 5820
6078 AACGGCTATATACGTTGCTTGTGCGGGGTAAACCAAGTCTCCCAACGCACTATGTGCC 6137
5821 GAGTGGATGCGTGGCAGCGGTGACCAACTACTTGTGCTTCCCTTACATTAACCAAGCTG 5880
6138 GAGAGGAGCGCGCGCGGGGTGTACTAGATCTCTCAAGCTTACATCACTCACTGAGTGG 6197
5881 CTCAGAAAGCTCCAACTGATTAATGAGAGTCCCATCCATGCGCGGCTGTGG 5940
6198 CTGAAAGGCTTCAATGATGAGATTAATGAGGACTGCTCCACGCTTGTGTCGCGCTGTGG 6257
5941 CTCGCGATGTGTGGGACTGGGTTTGCACATCTTAACAGACTTTAAATTTGGCTGAGCC 6000
6258 CTAAAGATGTTTGGACTGAGATATACGAGTGTATGATGACTTCAAGACTTGGCTTCAG 6317
5881 CTCAGAAAGCTCCAACTGATTAATGAGAGTCCCATCCATGCGCGGCTGTGG 5940
6198 CTGAAAGGCTTCAATGATGAGATTAATGAGGACTGCTCCACGCTTGTGTCGCGCTGTGG 6257
5941 CTCGCGATGTGTGGGACTGGGTTTGCACATCTTAACAGACTTTAAATTTGGCTGAGCC 6000
6258 CTAAAGATGTTTGGACTGAGATATACGAGTGTATGATGACTTCAAGACTTGGCTTCAG 6317
6001 TCCAAATTAATCCCAAGATAGCCGCGCTCCCTTGTGTCTGTCAAAAGGGGTCAAG 6060
6318 TCCAAAGCTCTGCGCGGGGTATCCGGAATCTCTTCTGTCAATGCCAAGCGGGGTCAAG 6377
6061 GCGGTGTGGCGGCACTGGCATCATGACCAAGGTGTCTTGGCGCGCAATATCTCT 6120
6378 GGAAGTCTGGGCGGGGATGGCATCATGCAAAACCACTGCGCATGTGGAGCACAGATCACCC 6437
6121 GCGAATGTCCGCTTGGGCTCCATGAGAAATCAAGGGGCTTAAGACCTGCATGAATATCTGG 6180
6438 GGAATGTCAAAATATGCTCTCATGAGATTTGTTGGGCAAAACCTTGCAGACACAGTGG 6497
6181 CAGGGGACTTCTCTTCAATGTTTACAGGAGGCGCAGTGTGCTGCGGAAACCCGCGCA 6240
6498 CATGGAACATTTCCCATCAACGATACCAAGCGGCGCCGTGACGCGCTCCCAAGCGCG 6557
6241 AACTTTAAGTCCCATCTGAGAGGTGGCGGCTCAGAGTACCGGAGGTGATGCGAGCAC 6300
6558 AACTATTCAGAGGCGCTGTGGCGGGGTGGCTGAGAGAGTACCTGTGAGGTTAAGCGGGGTG 6617
6301 GGGTCATACCATCAATACATACAGGACTCACACATGATTAATGGAAGTCCCTGCGCAACTA 6360
6618 GGGGATTTTCCATACGTAAGAGGAGATGACCATGACAAAGTGAATGCGCATGCAAGTT 6677
6361 CCTCTTCCGAGTTCTTTTCTGGGTGAGCGAGTGCAGATCCATATGAGTTTGCCTCCACA 6420
6678 CAGCGCCCTGAAATTTTTCAGAGAGTGAATGAGATACGTTTGAACAGATATGCTCCAGTG 6737
6421 CCGAAGCGTTTTTCCGGGATGAGGTCTCGTTCTGCGTTGGGCTTAATTCATTTGTGCTG 6480

```

Db      6738  TGCAAACTCTCTTACAGAGAGAGTGTGATTTCCAGGTCGGGCTCAACAGTACTGTGTC 6797
Qy      6481  GGGTCCAGCTTCTTTCGACACCTCTGAACCCGACACAGCTATTGATGTCCATCTACA 6540
Db      6798  GGGTCACAGCTCCCATGTGAGCCCGAACCGGATGTGAGTGCATCTCTCATCTCAC 6857
Qy      6541  GATCATCTCATATACAGCGCGAGACTGCAGCGCGGCTTTAGCGCGGGGTCAACCCCA 6600
Db      6658  GAACCTCTCATATTACAGAGAGACCGGCCAAGGTAGCTGTGGCAGGGGGTCTCCCCC 6917
Qy      6601  TCCGAGCGAAGCTCTTCGGCGAGCCAGCTATCGGCACCATCTGCGAGCCACTGCAAC 6660
Db      6918  TCTTTGGCCAGCTCTTCAAGCTTACGACAGTGTCTGCGCTTCTTTGAAGGCGACATGACT 6977
Qy      6661  ACCCAGCGCAAGCTTATGATGTGACATGTGTGATCTAACTT-----GTTCC 6708
Db      6978  ACCCATCATGACTCTCCCGAGCGTGAACCTCATCGAGGCGCAACTCTGTGGCGGACAGAG 7037
Qy      6709  ATGGGGGGCGATGTGACTCGGATAGAGTCTGGGTCCAAAGTGTGTGTTCTGGACTCTCTC 6768
Db      7038  ATGGGCGGGAACATCATCCCTGTGTGAGTCAAGAAATTAAGTGTATCTCTGAGCTCTTTC 7097
Qy      6769  GACCCAAATGATCGAAGAAAGAGCGACTTGAGCTTTCGATACCATGAGATACATGCTC 6828
Db      7098  GATCCGATTCGGGCGGTGAGAGATGAGAGGGAATATCCGTCCGCGGAGATCTCGGA 7157
Qy      6829  CCCAAGAAAGATTTCCACACAGCTTTTACCGGCTTGGGCGACGGCTGATTACAAACCCACG 6888
Db      7158  AAACCCAGGAAGTTCCTCCCGACGCTTCCCATATGGGACAGCCCGGATTTACMACCTCTCA 7217
Qy      6889  CTTGTGGAATCGTGAAGAAAGGCGAGATTACCAACGGGCACTGTTGGCGGCTGTCTCTC 6948
Db      7218  CTGTGTAGATCTGTGAAGAGACCGGACTAGTCTCCCGGTGTGTACACAGGATGCTTTTG 7277
Qy      6949  CTTCTCTCTTGAAGAAACCCCGACGCTTCCCCCAAGAGAGCGCCGAGACATGTGGGCTAAGT 7008
Db      7278  CCAATCTACAAAGGCCCCCGCAATACACCTTCAAGGAGAAAGAGAGCGTTGTCTGACA 7337
Qy      7009  GAGGATCTCATAGAGATGCTCTTCAACAGCTGGCCATTAACTCTTTGGCAGACCCCCC 7068
Db      7338  GAGTCCACCGTGTCTTCTGTGGCGAGCTGCTCTTAAGACCTTTGGGAGCTCCGGG 7397
Qy      7069  CCAAGCGGCAATTCAGGCTTTTCCAGGGGGCGGGCGCTGCGATTCGGGAGTCAAGAC 7128
Db      7398  TCTGTGCGCTTTGACAGCGG---CAGGCGACATGGCCCTCCG-----TCAAGCC 7445
Qy      7129  CTTCTCTGATGATGTGGCTCTTTTCGAGACAGGTTTCATCTTTCATGCCCCCTCTGAG 7188
Db      7446  TCCGACGACGGCGCAAAAGGATCCGAGTGTGAGTGTGATCTCTCATGCCCCCTCTGAG 7505
Qy      7189  GGGGAGCTTGGAGATCCAGACCTTGAGACCTGAGCAGGTAGAGCCCAACCCCCCGACG 7248
Db      7506  GGAGAGCCAGGGGAGCCCGACT-----7528
Qy      7249  GGGGGGGTGGAGCTCCCGGCTCGGAGCTCGGGGCTGTCTACTTCTCCGAGAGAGAC 7308
Db      7529  -----CAGCGACGGGTCTTGTCTTACCTGTGACGGGGAAGCTGT 7568
Qy      7309  GACTCCGTGTGTGTCTCTCATGTCTACTCTCTGACCGGGGCTCTTAATACTCTTGT 7368
Db      7569  GAGGAGCTGTGTGTCTCTCAATGTCTTAATACAGAGGTGTCTTTGATCAGCCATGC 7628
Qy      7369  AGTCCCGAAGAGAGAGATACCGATTAACTCTTTGAGCACTCTCTGTGGCATATAC 7428
Db      7629  GCTGCGGAGAGAGCAAGTTGCCATCATCTCGTTGAGCAACTTTTGTGCGTCAACAC 7688
Qy      7429  AACAAAGTGTACTGTACACAAAGAGCGCTCATTAAAGGGCTTAAGAGTAACTTTT 7488
Db      7689  AGATGTGTACTCTCAACATCTCTCGAGGCAAGTCTTGGCGAGAAAGAGTCACTTT 7748
Qy      7489  GATGAGTGAAGTGTCTGACTCTTACTAGACTCACTCTTAAAGGACATTAACTGACG 7548

```

```

Db      7749  GACAGACTGCAAGTCTGAGCAGCACTACCGGGACGTGTCTAAGAGATGAAGCGAAG 7808
Qy      7549  GCGTCCAAAGTCAACCGGAAGCTCTCTCAACATGGAAGAGGCTTGGCAGTTAAACCCACCC 7608
Db      7809  GCGTCCAAAGTCAAGGCTTGTATCTATAGAGAGGCTGTCAACATGAAGCCCTCCA 7868
Qy      7609  CATTCTCAAGATCTTAATATGGGTTTGGGCTTAAGAGGTTCCGACCTGTCCGGAGG 7668
Db      7869  CATTGCGCAAAATCCAAATTTGGGTACGGGCGAAGGACGTCCGAGACTTATCCAGAGG 7928
Qy      7669  GCCGTTAACCAATCAAGTCCGTGTGAGAGAGCTCTGAGAGACTGAGAAACCAAT 7728
Db      7929  GCGGTCAACCAATCCGCTCGTGTGGAGAGACTTGTGAGAGACATGAAACCAAT 7988
Qy      7729  CCACAAACCAATATGCGCAAAATAGAGTGTCTGCGTGAACCCACCAAGGGGGGAG 7788
Db      7989  GATGACACCAATGAGAAATAGAGTGTCTGCTCCAACTGAGAAAGAGGCGGC 8048
Qy      7789  AAAGCAGCTGCGCTTATCGTTTACCTGACCTCGGCGTCAAGGCTGCGAAGATGGCC 7848
Db      8049  AAGCAGCTGCGCTTATCGTATTCAGACCTGGGGTACGTGTATCGAGAAATGGCC 8108
Qy      7849  CTTTATGACATTAACAAAACTTCTCAGGCGGTGATGGGCGCTTCTTAATGATTCAG 7908
Db      8109  CTTTACGACGTGTCTCAACCTTCTCAGGCGGTGATGGGCGCTCATACGGAATTCAG 8168
Qy      7909  TATTTCCCGCTCAGCGGGTATGAGTTTCTTTGAAAGCATGGGCGAAGAAAGAGACCT 7968
Db      8169  TACTCTCTGGGACGCGGTATGAGTCTGTGAAATCTGGAATCAAGAAATGCTCT 8228
Qy      7969  ATGGGTTTTCTGATGATACCCGATGCTTGACTCAACCTGCACTGAGAGACATGAG 8028
Db      8229  ATGGGCTTCTCATATGACACCCCTGCTTGACTCAACGCTCACTAGGAATGACATCGT 8288
Qy      8029  ACTGAGAGTCTCATATATGCGGCTGTCTTTCGCGAGAGGCGCCACACTCCATACAC 8088
Db      8289  ACTGAGGAATCAATTTACATGTTGATGGTCCCGGACAGGCGAGCCATTAAG 8348
Qy      8089  TCGCTAACTGAGAGCTTTACGTGGAGGCTATGTTCAACAGCAAGGCGCAACCTGC 8148
Db      8349  TCGCTACAGAGCGGCTTATATGTGGGAGTCCCTGATTAATTCGAAAGGCGAGAACTGC 8408
Qy      8149  GGGTACAGGCTTTCGCGCGCAGCGGGTCTCAACCTGACATGAGGAAACACATACA 8208
Db      8409  GGTATATGCGGTGTGCGCGAAGTGGCTCTACATGCTGCGGCAACACCTTACA 8468
Qy      8209  TGTCTAGTGAAGCTTAAAGCTGTGAAGCTGCAAGGATTAATGCGCCACATGCTG 8268
Db      8469  TGTTAATTGAAGGCTACGTGGCTGTGAGCTGCAAAAGCTCCAGACTCAGATGCTC 8528
Qy      8269  GTATGCGGCGATGACTTGTGTCTATCTAGAAAGCAGGGAGCCAGAGAGAGAGAGCGG 8328
Db      8529  GTGAAGGAGACGACCTTGTGTATCTGTGAAGAGTGGGAAACCCAGAGAGATGCGCGG 8588
Qy      8329  AACCTGAGAGCTTTCACGAGGCTATGACAGGATTTCTGCTCTCTGCTGACCCCCC 8388
Db      8589  GCGCTTACAGCTTTCACGAGGCTATGATGATATTCGCGCCCCCGGAGACCGGCC 8648
Qy      8389  AGACCGGAGTATGATCTGAGCTGATTAACATCTTGTCTCTCAATGTGTGTGTGCGCTG 8448
Db      8649  CAACAGAAATACGACTTGAAGTGTATACGTATGCTCTCCATGTGTGTGCGTGCAC 8708
Qy      8449  GCGCCACAAAGCCCGCGAGATCTATCTGACCAAGAGACCTTACACTCCAAATGCGCGG 8508
Db      8709  GATGCAATCGGCAAAAGGTATCTATCTACCTCAGGAGCCGAGACCCACACCTCTGACGG 8768
Qy      8509  GCTGCTGGGAAACAGTTAGACACTCCCTGTCAATTCAATGCTGGGAAACATATCAG 8568
Db      8769  GCTGCGTGGAGACAGTTAGACACTCAGTCACTCTGCTGAGGCAATATCATCATG 8828
Qy      8569  TACGCCCCGACATATAGGCTGCGATGTGTCTGATGACACATTTCTTCATTTCTCATG 8628
Db      8829  TATGCGCCACCTTATGGGAGAGATGATGTGATGATCTATTTCTTCTATCTCTTCTA 8888

```

OY	86529	GCTCAAGACACGGCTGGACAGAAACCTCAACTTGTAGATGTAACGAGCGGTGACTCCGTG	8668
Db	8889	GCTCAGAGGCACTTGAAGAAAAGCCCTGATTTGTCAATCTACGGGGCCCTGTACTTCATTT	8948
OY	8689	AGTCCTTGGACCTCCAGCTATAATTGAAGTTACATG93CTTGAACGCTTTTCTCTG	8748
Db	8949	GAGCCACTTGAACCTTACCTCAGATCATTTGAAGACCTCCATG9TCTTAGCGCAATTTTCACTC	9008
OY	8749	CACACATCACTCCCAACAGAACTGACAGGGGTGGCTTAGGCCCTCAGAAAACTTGGGGCG	8808
Db	9009	CACAGTTACTCTCAGGTGAGATCAATAGGGTGGCTTAGCCCTCAGAAACTTGGGGTA	9068
OY	8809	CCACCCCTCAGAGCGCTGGAGAAGCCGGGCAAGTCAGTCAAG93GCTCCCTATCTCCGT	8868
Db	9069	CCGCTTTTGGCAGTCTGGAGACATCGGGCCAGAAGTGTCCGGGTAAAGCTACTGTCCAG	9128
OY	8869	GGGGGGAGAGCGGCGCTTTGGCGTGTGATATCTTTCAATTGGGCGGTGAACCAACGTC	8928
Db	9129	GGGGGAGAGGCTGGCCACTTGGCGGCAAGTACTCTTCAACTGGGGCAGTAAAGACCAAGTT	9188
OY	8929	AAACTCATCTCAATTGCGGAAAGCGGCGCTCCCTGATATTATCAAGCTGGTTGACCGTCGC	8988
Db	9189	AAACTCATCTCAATCCCGGCTGGCGTCCAGCTAAGACTTGTCCGGCTGGTTGTGTCTGT	9248
OY	8989	GCCGGCGGGGCGACATTTATATCACAGCGTGTGCGGTCCGACCCCGCTTATTTGCTCTTT	9048
Db	9249	TACAACGGGGAGACATATATCACAGCGCTGTCTCGTGCACCCCGCTGGTTATGTGTG	9308
OY	9049	GGCCACTCTCACTTTTGTAGGGGTAGACCTTTTCCAACTCCCGCGCTCGG	9099
Db	9309	TGCTCACTCTCACTTCTGTAGGGGTAGACCTTACTACGCTGTCCCAACCGG	9359

Search completed: October 31, 2005, 16:23:32
Job time : 3057 secs

QY 856 CTCGACGCCCAATGTCATTTCTCGCCGACGACCACTGTTTGTCCAAAGCTGCAAT 915
DB 412 CTTCGACGTTAGCTGATCA---TCTGCTCTCAGACACATAGATTGTGATGATGCAAC 356
QY 916 TGTCTCATCTACCTGTTGATCCTGATCCTGATGACACCGCATG-GCATGGACATGATGATGA 974
DB 355 TGTCTCATCTATCTCTGCGCCATCATCGACACCGTATGAGCATGGACATGATGATGA 296
QY 975 CTGGTGGCCACGCGCTACATGATCTTGGGTACGGGATGCTGTCCCGAGGCTCATTTAT 1034
DB 295 CTGGTCTGTGACCGCTGCTATGATCATGGGATGAGCATGGGCTTCTGAGGTATCAT 236
QY 1035 AGACATCATTTAGCGGGCTCATTTGGGCGCTCATGTTTGGCTTACTTCTTATGCA 1094
DB 235 AGATATCATCAGCGGGGCTCAGTGGGGCGTCAATGTTGGCTTACCTTCTATGCA 176
QY 1095 GGGAGCGTGGGCGCAAAATGCTGTCTCATCTTCTTGTGGCGCGGGGTGACCGCGGCAC 1154
DB 175 GGGAGCGTGGGCGCAAAATGCTGTCTCATCTTCTTGTGACCGCTGGGGTACACCGGACAC 116
QY 1155 CCATCTGTTGGGGGTTCTGTC-CCGCGACACACCGGGCGCTCACACGCTTATTTGACA 1213
DB 115 CCACTGACAGGTGGACCTGCTAGCGCGCGCGCGACGATTTGCTCAGCTTCTCATCC 56
QY 1214 TGGGCCCCAGCAGAAATCCAGCTCGTTAACAACCATGGCAGCTGG 1260
DB 55 CTGGCGCTAAACAGAACATCACTCATTTAGATTCATTTGCAATGG 9

RESULT 2
LOCUS AV758366 492 bp mRNA linear EST 19-OCT-2000
DEFINITION AV758366 BM Homo sapiens cDNA clone BMAKA03 5', mRNA sequence.
ACCESSION AV758366
VERSION AV758366.1 GI:10916214
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
COMMENT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 492)
AUTHORS Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H., Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,D., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
TITLE Homo sapiens cDNA BM clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@cnhg.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1..492
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMAKA03"
/issue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_id="BM"
/note="Vector: pTRIP1X2; Site_1: sflIA; Site_2: sflIB"

ORIGIN
Query Match 2.6%; Score 235.8; DB 1; Length 492;
Base Local Similarity 72.0%; Pred. No. 1,7e-53;
Matches 347; Conservative 0; Mismatches 130; Indels 5; Gaps 3;

QY 783 CATGACATGATGATGATGTCGCGACGCTGCTCTGCTCTGCTCTGATGAGGAGACCTCTG 842
DB 486 CACCTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 427
QY 843 GGTGGGGTGTATGCTTCGACGCCCAATGTTCTATGTTCTGCGCGGACCACTGTTTGT 902
DB 426 CTTTGGGGTATGCTTGGACGCCCACTG---ATTATCTCTGACGACCAATTTGGTTGT 370
QY 903 CCAAGACTGCAATGCTCATCTACCTGATGATGATGATGATGATGATGATGATGATGATG 961
DB 369 GCAAGATGCAATGCTCATCTATTTATCTGCGCTGATCACTGATCACTGATGATGATG 310
QY 962 ACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1021
DB 309 CTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 250
QY 1022 CCGAGCTATTTAATGATCATTTAGCGGGCTCATTTGGGGGCTCATTTTGGCTTGGCT 1081
DB 249 CTGAAGTCTCTATGATATGATGATGATGATGATGATGATGATGATGATGATGATG 190
QY 1082 ACTTCTATGACAGGAGGATGAGGAGAAAGTGTGATCATCTTGTGCGCGCGGAG 1141
DB 189 ACTTCAATGACAGGAGGATGAGGAGAAAGTGTGATCATCTTGTGATGATGATGATG 130
QY 1142 TGGACGCGGACCCCATCTGTTGGGGGTTCTGCGCGGACGACGACGCGGCGCTCACCA 1201
DB 129 TTGATGCGGACCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 70
QY 1202 GCTTATTTGA-CATGGGCCCCAGCAGAAATCCAGCTCGTTAACAACCATGGCAGCTGG 1260
DB 69 GGTTCATGATGCTTGGGCTTAAAGATGATGATGATGATGATGATGATGATGATGATG 10
QY 1261 CA 1262
DB 9 NA 8

RESULT 3
LOCUS R28798 179 bp mRNA linear EST 03-DEC-1999
DEFINITION R28798 F0-266D 22 week old human fetal liver cDNA library Homo sapiens
ACCESSION R28798
VERSION R28798.1 GI:6514170
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
COMMENT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 179)
AUTHORS Choi,S.-S., Yun,J.W., Choi,E.K., Cho,Y.G., Sung,Y.C. and Shin,H.S.
TITLE Construction of a gene expression profile of a human fetal liver by single-pass cDNA sequencing
JOURNAL Mamm. Genome 6 (9), 653-657 (1995)
MEDLINE 96081342
PubMed 8535075
COMMENT Contact: Hee-Sup Shin
Developmental Genetics
Pohang Institute of Science & Technology
San31, Hyojadong Pohang, 790-784 Republic of Korea
Tel: 562-279-2291
Fax: 562-279-2199
Email: shins@vision.postech.ac.kr
Seq primer: T3 primer.
Location/Qualifiers

FEATURES
source
1..179
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="F0-266D"
/lab_host="XL1-blue MRF"
/clone_id="22 week old human fetal liver cDNA library"

/note="Vector: pBluescriptII SK(-); Site 1: EcoRI; Site 2: XhoI; The cDNA library made by oligo-dT primed and directionally cloned between 5'XhoI 1' sites."

ORIGIN

Query Match 0.6%; Score 57.6; DB 7; Length 179;
Best Local Similarity 78.4%; Pred. No. 0.0056;
Matches 69; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1210 GACATGGGCCCCGAGGAGAAATCCAGCTGTGTAACACCAATGGAGCTGGACATCAAC 1269
DB 50 GACATTGTAGATCCGAGAACATTCCTGTAAGCCTAATGCGAGCTGGACATCAAC 109

QY 1270 CGCAGCCGCTGAACTGCAATGACTCT 1297
DB 110 AGGACTGCTTGAATTGCAATGACTCT 137

RESULT 4
LOCUS CN373175 522 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000531390153 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN373175
VERSION CN373175.1 GI:47373109
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 522)

AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebowitz, J. and Stanton, L.W.

TITLE Transcriptional characterization elucidates signaling networks that
control human ES cell growth and differentiation

JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)

COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert length: 522 Std Error: 0.00.
Location/Qualifiers

FEATURES
source 1..522
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/issue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
/note="oligo dT primed, full-length enriched cDNA library
from embryoid body outgrowth derived from hES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

ORIGIN

Query Match 0.6%; Score 52; DB 7; Length 522;
Best Local Similarity 46.8%; Pred. No. 0.029;
Matches 163; Conservative 0; Mismatches 185; Indels 0; Gaps 0;

QY 748 CGGCCCCGGGCTCTACGAGGCTGGCGAGCAATGATGTTGATGATGCCGC 807
DB 59 CAGCCCAATGTCAGCCCTGTCAGCAGCACTTGGGCAATGTGCGCTCATGAG 118

QY 808 AGCGTGTGCTGCGCTTACGTCGAGGAGCTTCGCGTGGGTGATGTCGAGCCCA 867
DB 119 CTGAGGCGCTTTCCTCCAGGAGAGTGGGGGCGCGGCGAGGCTGACCCCGGTATAC 178

QY 868 ATGTTATGTTCTCGCGCAGACCACTGTTGTCGAAGTGCATTTGCTCATTCAC 927
DB 179 CCTGACAGCGCTTGTGCTGCTCTCTTTCGCCACATCATCTACATCTCAAC 238

QY 928 CCGTACCATCAGTACGACCGCATGGCATGGAGACATGATGTAAGTGTGCCACG 987
DB 239 CACAGCTTCATCCGTGTGTCCCGAAGAGCTGACATCTGTGAATCTTGCTTCCAC 298

QY 988 GCTACCATGATCTTGGCGTACGCGATGCTGTCCCGAGTATATAGATCATTAAC 1047
DB 299 ATAGCATATACCTCTGTCTTTTGGGGGCGATACACTACCAACATCAGATGTC 358

QY 1048 GGGCTCATTTGGGCGCTCATGTTGGCTTGCTTCTTATGAG 1095
DB 359 TCCAGAGCGGTGGGATCACCTGACATCTCTCTCATCACGCTG 406

RESULT 5
LOCUS CNS0091P 925 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TERT end of BAC #
BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL053013.1 GI:4934461
VERSION AL053013.1
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 925)

REFERENCE Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT

Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutopo Oosawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source 1..925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone_lib="BACPAC19D16"
/clone_lib="RPCI-98"
/note="end : TERT3"

ORIGIN

Query Match 0.6%; Score 52; DB 9; Length 925;
Best Local Similarity 12.8%; Pred. No. 0.035;
Matches 46; Conservative 167; Mismatches 147; Indels 0; Gaps 0;

QY 668 TCCCGGAGTGTCCCGTCGAGAAAGTGGAGATCATCTAGTCTGATACCGGCT 727
DB 554 TSSGGGAGGCGSSGGBSCSSCSBCCBCCCGCCSSYCCSSSSSSSSSCST 613

QY 728 CACCAATGTGCGCGGAGCGCGCGCGCTTACGAGGCTTGGCGAGCATATCG 787
DB 614 SBSGCGCCSKSVGCTGSSSSSSSSSSSSSTSSSTSSSSSSSSSSSSSTTAK 673

QY 788 ACATGTTGATGTCGCGCAGCTGTGCTTACGTTGGGAGCTCTGCGGCTG 847

Db 674 STSAGSGSWAGGSGTSTSSSSSSTSTSSSVSSGSKSSTBSAGSBSSGSSSS 733
Qy 848 GGGGATGCTCGAGCCCAATGTTCTATGTCGCGCAGCACCTGTTGTCCAG 907
Db 734 SSTSSBBSCTSTSSSSSSSTSSCTCCCTCCSYSSSTSSSSSTSSWGTSSSSSSVG 793
Qy 908 ACTGCAATTCCTCATCTACCTGTGATCCATCACTGACACCCGATGCGATGAGATGA 967
Db 794 TSSSSDSTSTCCSCCYMCCTCCSYBMBCTSTSSCGSSSSSGKGVTKCGCGGSSST 853
Qy 968 TGATGAACGTGTGCGCCACCGCTACCATGATCTTTGGGTACGCGATGCTGCTCCGAG 1027
Db 854 NGMBGTSSACSS 913

RESULT 6
CNS0091P/c 925 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL053013
VERSION AL053013.1 GI:4934461
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 925)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamooser in Pleter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp. the same strain used for the BDGP's
P1 and BGT libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
SOURCE
1. 925
Location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_1lb="RPCI-98"
/note="end : TET3"

ORIGIN
Query Match 0.6%; Score 51.8; DB 9; Length 925;
Best Local Similarity 15.6%; Pred. No. 0.04;
Matches 73; Conservative 178; Mismatches 211; Indels 5; Gaps 1;

Qy 6335 CGGGCTGTGCTCTCTCTCTCTAGAAAACCCGAGCCTCCCAAGAGCGCGCGA 6394
Db 925 CSBSCSCSCSS 866
Qy 6995 CAGGAGGCTTAAGAGAGCTCATAGAGATGCTTCAACAGCTGCGCATTAAGCT 7054
Db 865 SSGTSSAKVCNMASSSSCGCCCGMAABCCMSCSSSSCCGSASARVKVRAAGAGAKRGGS 806
Qy 7055 TTGGCAGACCCGCCCAAGCGCGATTTCAGGCTTTCCACGCGGCGCGGCGCTGCGATT 7114

Db 805 GGAASASHSSSSACBSSSSSSSCASCSWASSSSSASBSRSGGAGCGSASSSSSSSSS 746
Qy 7115 CCGGACATCAGACGCTCTCTATAGTTGGCCCTTTCGAGACAGATTCTCTTCCA 7174
Db 745 AAGAGSVASASSSSSSSSSVSCSSVASMSSCSSSSSSASASSSSSSSSACACSCCC 686
Qy 7175 TGCCCCC-----CTCGAGAGGGAGCTTGAGATCAAGACTGAGAGCTGAGCAGATGA 7229
Db 685 TMSGCSCTASMASARSSSSSSSSSCSSSMASASSSSASASSSSSSSSSSSAGCB 626
Qy 7230 GCCCAACCCCCCCCCAGGGGGGTGCGAGCTCCCGCTCGACTGGGCTCTGTC 7289
Db 625 SMSGGGSGSVASAGMSGVSSVSGSGSGGCGGCGGSGSSGSGSGGSGVCGCS 566
Qy 7290 TACTGTCTCGAGAGAGCACTCCGTGTGTCGTCTCCATCATCTCCTGGACCGG 7349
Db 565 SGCMKRCSCSSAAAACVAAACCMPCGKSGCTNNNNNTNTNTNTNTNTNTNT 506
Qy 7350 GCGCTTAATTAATCTCTCTGTAGTCCGAAAGAGAGAGATTACCGATT 7396
Db 505 ATAATAANNNT 459

RESULT 7
CNS015Y4/c 1203 bp DNA linear GSS 26-JUL-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC
DEFINITION BACN15E10 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL106054
VERSION AL106054.1 GI:5619805
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1203)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP).
<http://www.edgp.ebi.ac.uk>. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES
SOURCE
1. 1203
Location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN15E10"
/clone_1lb="DrosBAC"
/plasmid="pBelobAC11"
/note="end : T7"

ORIGIN
Query Match 0.6%; Score 51.4; DB 9; Length 1203;
Best Local Similarity 34.6%; Pred. No. 0.057; Indels 2; Gaps 1;
Matches 128; Conservative 69; Mismatches 171; Indels 2; Gaps 1;

Qy 3322 GCTGAGGAGGACTTGTAGAGGTGCGCCAGCCCTCGGAGCTTAATCTTTGAGCCGTGC 3381
Db 1198 GCGGG 1139
Qy 3382 ACGTGTGAGCGGTGACCTGTACTGTACGCGGAACGCTGATGTATCCCGGCTCA 3441

Db	1138	CSGGGGGGCCCCGGGGGGCCCCCCCCSSSSGGSGGSSGGGGGGGGGGGGGGGGGGCCG	1079
Qy	3442	AGAGCGGGGGGAAACGGGGAGCGCTACTCTCCCGAACCCTTTTCCACTTTGAAGGG	3501
Db	1078	GGSGGGGGSSSSSSSSGG	1019
Qy	3502	TCCTCAGAGAGCCCGG--TGCTATGCGCCAGAGGGCAGCTGTGAGTCTTTCGGGCGAG	3559
Db	1018	CCGCSGG	959
Qy	3500	CTGTGTGTCTCGGGGGCGGTAGCTAAGTCAATGATTATCCCGCTTGAAGACTCGACA	3619
Db	958	CCCBTKCBKCCCKSSCCBKBCCCBBCSSCCSBKCCCCCNCCCCCCCCCCCCCCCCCS	899
Qy	3620	TGCTGACGCGGTGCCCACTTTAGTGAACAAGACACCACTGTGTGTGCCGAGACT	3679
Db	898	CCCCSSGGGSSCCCCCCCCGGSGGAGNCCCCCCCCCGGGGGGGGGGSCCSGGGGCC	839
Qy	3680	ATCAGTCCG	3689
Db	838	CCCCSSGG	829

RESULT	8			
HSM804270				
LOCUS	HSM804270	4762 bp	mRNA	linear
DEFINITION	Homo sapiens mRNA; CDNA DKFZp666M0710 (from clone DKFZp666M0710).			HTC 22-SEP-2004
ACCESSION	AL332959			
VERSION	AL332959.1	GI:21735547		
KEYWORDS	HTC.			
SOURCE	Homo sapiens			
ORGANISM	Homo sapiens (human)			

REFERENCE 1 (bases 1 to 4762)
AUTHORS Ansoer, W., Krieger, S., Regiert, T., Rithmueler, C., Schwager, B.,
Mewes, H. W., Well, B., Amd, C., Osanger, A., Fobo, G., Han, M. and
Wiemann, S.

CONSRNT TITLE	JOURNAL	COMMENT
The German cDNA Consortium Direct Submission		Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuharby, GERMANY
		Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by EMBL (European Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.
		This clone (DKFP666M0710) is available at the RZPD Deutsches Resourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFP666M0710 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/ .

FEATURES	Location/Qualifiers
source	1. .4762

```

gene
1..4762
/gene="DKFZp666M0710"
<2..2719
/gene="DKFZp666M0710"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH56283.1"

```

```

/db_xref="GI:52545734"
/translation="SHIGWASGWEECTTAAAGNASKKVEIVLETSATSCAERYA
NMNRDRFMRPTLIGTAYOSLCLOYRPFVSGGAPRSCDPAQREPDYSHC
LITDITRVLTYTFLPAINASNAITLHQRVYTLKAASISDMDDVYVAAQKFLG
YVDDIKELVEWMDMASNLMLVDBHLMLQREKACASRIVGLERIGGALSFMQH
ISVANRVALLAELIKPHSYVGLCTAPQRECGVPTGRSGQNPDPPEPPADQ
LRFECTTGVRNPTSGAAGPCKRGKAVPVPVFACTSGQGNLTVEPVALSRJHAAGEA
NGLRPFHSNSTPSRGAAGPCKRGKAVPVPVFACTSGQGNLTVEPVALSRJHAAGEA
PVALRMOEQEPGEAGCWTSBGCORRSQNPVALSHCHLGNLVALVIMELASIPFEVGA
GAGLHPVVPCTALLLCTPATIITYTILNMSIRVSKGKMLNLNCTCLFAMTASVA
GGILTNVQWCOAVGITHYKSLITLTMNGVARVHLKELTYRAPPOGCDALPFP
SPMLRFLIAGGTPILICGTTAIVNHNHYDHSYCPMLVRSIGAVIIVALLITIT
MIVYLCGLILRGLPAINONPKRNGSPASLXGZGCTLRPGRGGPILSDGSLILTS
ARVOTPEPPEPDGSLSPYQULKGAGHNLPUTGHVGLRSGSYPALAAGCYQLVLR
GGSLRPGPRLKLSLCOABGRSIVALLPESLCISRGPRKRGARRRFRFVPGG
APLQOVLPKROQSRASGSLQAHQPAAGESGSGGGGGRKGRGARGAGGHPCGPRP
PPQORARPARAGAOBPGQITRGGQLRQEPAQGRGARGGAGGAAYORER"

```

ORIGIN

Query Match	0.64;	Score	50.6;	DB	3;	Length	4762;
Best Local Similarity	42.64;	Pred. No.	0.15;				
Matches	263;	Conservative	0;	Mismatches	354;	Indels	0;
				Gaps			0;

LOCUS	DEFINITION	AY411015	3661 bp	DNA	linear	GSS 16-DEC-2003
OY		684	GTGGAGAAAGTGGGGAATGATCTCAGTCTGGATACCGGTCTCACCGAATGAGCCGT			743
Db		1195	GAGCCAGAGAGGAGCCCGGAGAGGTGGAGGCTGGACCTCGAAGGCGTGCAGCTCCGCTC			1254
OY		744	GCAGCGGCGCCGCCCTCAACGACGAGGCTTGGCGGACGACATCGACATGGTTGTGATGTC			803
Db		1255	CAGCCAGGCCAATGTCAAGCGCCCTGCACTCGCAGCACTTGGGCAATTGGCCGGTCTAT			1314
OY		804	CGCCACGCTGTGCTCTGACCCTTAACGTGGGGGAACTTGGGGTGGGTATGCTCGAGC			863
Db		1315	GGAAGTAAAGGCTTTTCCAGAGGAGTGGGAGGCGCGGAGGAGGCTGACCCCGGT			1374
OY		864	CCAATATTTCAATTGTCTCGCCGACGACCACTGGTTTGTCAAGACTGCAATTGCTTCAT			923
Db		1375	ATACCCCTGCACGAGCCCTGTGCTGTGCTCTGCTCTTTCGCAACATATCACTAATCCT			1434
OY		924	CTACCTTGATCACTATCACTGGAACACCGCATGAGCATGATGATGAATGTGTGCC			983
Db		1435	CAACCAACAGCTCCATCCGTGTGTGCCGAAAGGCTGGCAATGCTGTGAACCTTGTCTT			1494
OY		984	CAGCGCTACCATGATCTTGGCGTACGCGAATGCGGTCCCGAAGTCACTTATAGCATCAT			1043
Db		1495	CCACATAGCCATGACCTCTGTGTCTTGGGAGGAGATACACTACCACTACCACTACAGAT			1554
OY		1044	TAGCGGAGCTCATTTGGGCGTCAATGTTGCGCTTGGCTACTTCTTATATGACGAGAGCGTG			1103
Db		1555	GGTGTGCAAGCGGCTGGGCAATCAACCCGCACTACTCTCCCTATTCACGCTGTGCTGGAT			1614
OY		1104	GGCGAAAGTGTGTATCTCTTCTGTGGCCCGCGGAGGTGAAGCGCGCACCCATACTGT			1163
Db		1615	GGGGCTGAAGGCGGAGATGCTCCATTAAGAGCTCACCTGGAGGAGCACCCCTCCGCAAGA			1674
OY		1164	TGGGGGTTCTTGGCGCGGACACACCGGAGCGCTCACAGCTTAATTGACATGGAGCCCGAG			1223
Db		1675	AGGGGAGCCCGCTTGTGCTTACTCCCACTCTTAATGCTCGGTTCTATTGATTCGTGAGG			1734
OY		1224	GCAGAAATCCAGCTCGTTAAACACCAATGACGTGGCACTGAACATCAACCGCACGCGCTGAA			1283
Db		1735	GATTCGACTATTAATCTGTGGCATCAACAGCTGACAGTCAATCATCAACAACTACGAGGACCA			1794
OY		1284	CTGCATGATCTCTTTCG 1300			
Db		1795	CAGCCCTTACTGTGTC 1811			

ACCESSION	AY111015
VERSION	AY111015.1 GI:39766983
KEYWORDS	GSS.
ORGANISM	Human sapiens (human)
REFERENCE	Clark,A.G., Gianowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
AUTHORS	Clark,A.G., Gianowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL	Science 302 (5652), 1960-1963 (2003)
PUBMED	14671302
REFERENCE	2 (bases 1 to 3661)
AUTHORS	Clark,A.G., Gianowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
TITLE	Direct Submission
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES	Location/Qualifiers
source	1..3661
gene	/organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" "1..>3661 /gene="GPR124" /locus_tag="HCM4083"
ORIGIN	
Query Match	0.5%; Score 50; DB 9; Length 3661;
Best Local Similarity	16.7%; Pzed. No.0.2; Matches 170; Conservative 0; Mismatches 851; Indels 0; Gaps 0;
Db	75 GGGCGCGCGCGAGATCGTTGCGGAGATACTTGTTCGCGCGAGGCGCCAGGTTGG 134
Qy	1264 GGCCTGCGACCGCATCTGTGGTGGCTCTGAGCGCATTTGGGGGGCGCCCTTAGCCCCCA 1322
Db	135 TGTGCGCGCGACAAAGAACTTCGAGCGCGTCCAGCGCATGTGAAGCGCGCACCGCAT 194
Qy	1324 TGCCAGACACCTCAGTGNNN 1383
Db	195 CCCTTAAAGATCGGCGCTCCACTGTGCAATCTCTGGGAGAAAACGAGATACCCCTGCGCT 254
Qy	1384 NNN 1443
Db	255 ATACGGGAATGAGGAGACTCGGCTGGGAGAGATGGCTCCTGTCCCCGAGATTCCCGTCC 314
Qy	1444 NNN 1503
Db	315 CTCTTGGGCGCCCAATGACCCCGGATAGTGTGGGCAACGTGGGTAAGTCATGATAC 374
Qy	1504 NNN 1563
Db	375 CCTAACGTGGCGGCTTTCGCGACCTCATGGGATACCTCTGTGTGGGCGCCGCGTGG 434
Qy	1564 NNN 1623
Db	435 CGGCGTCCGCAAGCTCTCGCGCATGGGCTGAGAGTCTCTGAGAGACGGGGTTAATTTCG 494
Qy	495 AACAGGAACTTACCCGCGTGTCTCTTTCATCTTCTTGTGCGCTGTGTCTGAT 554
Db	1664 NNN 1743
Qy	555 CACCAACCCGCGTCTCGCTGCGGAAAGTGAAGACATCAGTACCGGCTACATGTTGACTTA 614

[illegible]

Class: transposon-tagged.
Location/Qualifiers
1. .393
/organism="Zea mays"
/mol_type="genomic DNA"
/cultiivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="1007 - RescueMu Grid H"
/note="Organ: leaf; Vector: RescueMu (engineered from pBluescript backbone); Site_1: BamHI; Site_2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu'. Grid H was grown at Berkeley in 2001. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

ORIGIN

Query Match 0.5%; Score 47.2; DB 8; Length 393;
Best Local Similarity 47.1%; Pred. No. 0.58;
Matches 145; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 1621 TTGAACGACGACTGACACCGCTGGGTCATGGTGGCTGCACGTGGATGAATCTTCT 1680
DB 334 TTGTGCTCCGCGACCGTGGCGCGCGCTTCGGCAGCGCGACGACGACTTAATG 275
QY 1681 GGCTACACCAAGACTTSCGGCGACACACCTGCGCTACTAGAGTGACTTCAACGCGACG 1740
DB 274 AGCTCCCGCAGACCTTCTGCTCCACCCGACCGCGCGCTCCGCGACGACGACGCGC 215
QY 1741 ACAGACTGTTGGCCCCCAGCACTGTTTGAAGAGCATCTGATACCTTAACCTTCAA 1800
DB 214 AGGTCAGGCGCTACACACACACACATCTGCTGCTGCTCTCTCTCTCTCTCTCT 155
QY 1801 TGGGCTCTGGGCGCTGACCGCGCAAGTGGCTGATGACCTTACCGCTACAGCTCTG 1860
DB 154 CGCGCGACGACGCGCGACCGCGCGCGCTGACGACCGCGCTGACGCGCTGACGCG 95
QY 1861 CATTACCCCTGCAAGTTAATCAATCATCTTCAATAAGATGATGTGGAGGGGTT 1920
DB 94 CCAGAGTCTGCAAGCGCGGTGTAAACAATATGATGATCTCTTCTCTCTCTCTAGAC 35
QY 1921 GAGCAGC 1928
DB 34 TAGCAAG 27

RESULT 13
CNS010EM 1009 bp DNA linear GSS 26-JUL-1999
LOCUS Drosophila melanogaster genome survey sequence SP6 end of BAC
DEFINITION BACN03P19 of DrosBAC library from Drosophila melanogaster (fruit fly). genomic survey sequence.
ACCESSION AL098882.1 GI:5610493
VERSION GSS.
SOURCE Drosophila melanogaster (fruit fly)
KEYWORDS Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1009)
AUTHORS Genoscope.
JOURNAL Direct Submission
TITLE Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr

COMMENT
- Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - <http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billard at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelBAC11.
FEATURES
source
1. 1009
Location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone_1lb="BACN03P19"
/plasmid="pBelBAC11"
/note="end : SP6"

ORIGIN

Query Match 0.5%; Score 46; DB 9; Length 1009;
Best Local Similarity 31.6%; Pred. No. 1.7;
Matches 99; Conservative 62; Mismatches 152; Indels 0; Gaps 0;

QY 145 ACAAGGAGACTTCGAGCGGTCCAGCCAGCTGGAAGCGCCAGCCATCCCTTAAGAT 204
DB 510 ACAACAAAGAAAYACACGACACGATGCBCTRTVACNCACVCMCCCMACAY 569
QY 205 CGCGCTCCACTGCGAATTCCTGGGGAAACGAGATACCCCTGACCCCTATACGGGAT 264
DB 570 SSCSCCCMCNCAAGNACARASRVACACMKVRAASAGRAVAMMACTAGSGCDSGGC 629
QY 265 GAGGACTCGCTGGGCGAGATGCTCTGTCCTGCTCCCGGAGTTCCCGTCTTGGGCG 324
DB 630 TRGGGSRSSGGGGGCGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 689
QY 325 CCAATGACCCCGGATGATGTCGCGCAAGTGGTGAAGTATGATACCTTAACGTCG 384
DB 690 CCCGCGBGSMSYGGCCGCGGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 749
QY 385 GGGCTTGCGGACCTATGGGCTACATCCCTGCTGCTGGGCGCGCGCGCGCGCGCG 444
DB 750 SCCSSGSCSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSG 809
QY 445 AGAGCTTCGCGC 457
DB 810 CCGGSCCGGSS 822

RESULT 14
CK307855/c 755 bp mRNA linear EST 01-MAR-2004
LOCUS SB0204482D10.fl normalized Keck-Tsag Library SB02 Taeniolopygla
DEFINITION guttata cDNA clone SB0204482D10.fl 5, mRNA sequence.
ACCESSION CK307855.1 GI:44817429
VERSION EST.
KEYWORDS Taeniolopygla guttata
SOURCE Taeniolopygla guttata
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauaria; Aves; Neognathae; Passeriformes; Estrilidae; Estrilidae; Taeniolopygla.
REFERENCE 1 (bases 1 to 755)
AUTHORS Clayton,D.F., Arnold,A.P., Ball,G.F., Brenowitz,E., George,J.M., Mello,C.V., Wade,J., Replogle,K., Lewin,H., Band,M., Hernandez,A. and Liu,L.
TITLE The Songbird Neurogenomics Initiative: An Evolving Public Resource for Study of Genes, Brain, and Behavior
JOURNAL Unpublished (2004)
COMMENT Contact: David F. Clayton
University of Illinois
B107 CStL, 601 S. Goodwin, Urbana, IL 61801, USA
Tel: 217 244 3668

Fax: 217 244 1648
Email: dclayton@uiuc.edu
Base Calling/Quality Scores: PHRED from Washington University
Genome Center.

Vector Trimming: Cross match from Washington University Genome
Center PHRAP suite. Low quality bases (Phred score < 20) were
trimmed from both ends of the sequence by an in-house script.
This sequence is vector free and at least 200 bp in length. Funded
by PHS grant # R01 NS045264, 'Songbird Neurogenomics Initiative.'

PCR Primers

FORWARD: TATACGACTCACTATAGG (T7)

BACKWARD: ATTACCTCACTAAG (T3)

Insert Length: 755 Std Error: 0.00

Plate: SB0204482 row: D column: 10

Seq primer: TATACGACTCACTATAGG (T7)

High quality sequence stop: 755.

FEATURES

source

1..755

/organism="Taeniopygia guttata"

/mol_type="mRNA"

/db_xref="taxon:59729"

/clone="SB0204482D10.f1"

/issue_type="Brain"

/dev_stage="late embryo", post-hatch days 1, 10, 20, 45,

and adult (pooled) "

/lab_host="DH10B"

/clone_lib="normalized Keck-Tagu Library SH02"

/note="Organ: Brain; Vector: pBS II SK(+); Site_1:

EcoRI(5' side of insert); Site_2: NotI (3' side of

insert); The library was constructed and normalized as

described by Bonaldo, M.F., Lennon, G. and Soares, W.B.

(1996), Genome Research 6(9): 791-806. An identifying tag

was added at the 3' during cDNA synthesis:

insertAAAAAAAAAAAAAAAAATCGCA."

ORIGIN

Query Match

Best Local Similarity 48.6%; Pred. No. 1.8; Length 755;

Matches 125; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 1166 GGGGTTTCGGCGGAGACACCGGGGCTTCAACGATTATTGACATGGGCCCAAGC 1225

DB 353 GGATTCTGCTCATAGCGCTCTTATTCTTCTCAAGCTTCTGCGGCTCTCCCA 294

QY 1226 AGAAATCCAGCTGTTAACAAGGACGTCGACATCAACCCGACCCCTGAATC 1285

DB 293 GGAATGCCCGGCTTCTCCGCGGACGCGCTCTTCAAGCCCTTGGCGCTCTC 234

QY 1286 GCATGACTCTTGCACACCGGCTTATCGCTCTGTTTACACCAAGCTTCAACT 1345

DB 233 CCAAGCTGCTTCTCCGCGGAGACGGGCTTCTTCTTCTCAAGCACAAGACCCGCT 174

QY 1346 CGTCAAGATGTCGGAACGATTCGCTGCGGACAGTATCGAGGCTTCCGGGTGGAT 1405

DB 173 TGGCAGCAAGTACTTGGCCGCTCTCGCGCGCTCTTGGCTTGTGCTTGCTT 114

QY 1406 GGGGCGCCCTTGCAATAT 1422

DB 113 GGGCTTCTGGGATCT 97

RESULT 15

CA733786

LOCUS CA733786 562 bp mRNA linear EST 26-NOV-2002

DEFINITION wipic.pk006.d15 wipic Triticum aestivum cDNA clone wipic.pk006.d15

ACCESSION CA733786

VERSION CA733786.1 GI:25549384

KEYWORDS EST

SOURCE Triticum aestivum (bread wheat)

ORGANISM Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

REFERENCE

Poideae; Triticaceae; Triticum.
1 (bases 1 to 562)
Tingey, S.V., Powell, M., Wolters, P., Dolan, M., Hainey, C., Yuan, Z.,

AUTHORS

Miao, G., Caraher, N. and Hanfey, M.K.

TITLE

Unpublished (2002)

JOURNAL

Unpublished (2002)

COMMENT

Contact: Scott V. Tingey

Crop Genetics

E.I. Dupont de Nemours and Company

1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA

Tel: 302-631-2602

Fax: 302-631-2607

Email: Scott.V.Tingey@USA.dupont.com

Seq primer: M13.

FEATURES

source

1..562

/organism="Triticum aestivum"

/mol_type="mRNA"

/db_xref="taxon:4565"

/clone="wipic.pk006.d15"

/issue_type="lemma and palea"

/lab_host="DH10B"

/note="Vector: pBluescript SK+; Site_1: EcoRI; Site_2:

XhoI; Wheat (Triticum aestivum, H1 line) lemma and palea"

ORIGIN

Query Match

Best Local Similarity 55.8%; Pred. No. 1.8; Length 562;

Matches 87; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 7167 CTTTTCAGAGCCCGCTGAGGGGAGCTTGAGATCGAGACTGGAGCTGAGCAGGT 7226

DB 318 CTCGCTGATGCGAGCTCGAGAGAGATTGCGCGGAGACTGAGCCCGGACCC 377

QY 7227 AGAGCCCAACCCCGCCAGGGGGGTGACAGCTCCGAGCTCGAGCTGGAGT 7286

DB 378 GAGAGCTCGGCTTCTGCTGAGGCGCTCGAGCAGAGCTCGGCTGCGCGGCGG 437

QY 7287 GTTACTTCTCCGAGAGAGAGACTTCGCTGTG 7322

DB 438 CCGGCGCTCTCTGTCGAGACGCGCGGGGCTG 473

Search completed: November 1, 2005, 04:05:44

Job time: 17327 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 31, 2005, 15:25:30 ; Search time 937 Seconds
(without alignments)
15889.522 Million cell updates/sec

Title: US-09-980-559-1_COPY_341_9439

Perfect score: 9099
Sequence: 1 atgcagcacaatccctaacc.....tttccctaccctccgcgcgcg 9099

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCITUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8494.2	93.4	9589	1	US-07-925-695-1 Sequence 1, Appli
2	8494.2	93.4	9589	1	US-07-925-695-2 Sequence 2, Appli
3	5653.8	62.1	9511	1	US-07-925-695-7 Sequence 7, Appli
4	5651	62.1	9511	1	US-07-925-695-6 Sequence 6, Appli
5	4141.8	45.5	9413	4	US-09-827-688-6 Sequence 6, Appli
6	4136	45.5	9599	3	US-09-014-416-6 Sequence 6, Appli
7	4129.6	45.4	12980	3	US-08-811-566-5 Sequence 5, Appli
8	4129.6	45.4	12980	3	US-09-034-756-5 Sequence 5, Appli
9	4128	45.4	9646	3	US-08-811-566-1 Sequence 1, Appli
10	4128	45.4	9646	3	US-09-034-756-1 Sequence 1, Appli
11	4121.6	45.3	9599	3	US-09-014-416-2 Sequence 2, Appli
12	4118	45.3	9401	5	PCT-US91-02225-9 Sequence 9, Appli
13	4115.2	45.2	9379	3	US-09-388-874-1 Sequence 1, Appli
14	4115.2	45.2	9379	4	US-09-916-359-1 Sequence 1, Appli
15	4115.2	45.2	9401	1	US-07-910-760-9 Sequence 9, Appli
16	4115.2	45.2	9401	1	US-08-440-519-9 Sequence 9, Appli
17	4115.2	45.2	9401	1	US-08-440-519-9 Sequence 9, Appli
18	4115.2	45.2	9401	3	US-08-823-895A-25 Sequence 25, Appli
19	4114.8	45.2	9379	3	US-08-444-818-176 Sequence 176, App
20	4112	45.2	9365	4	US-09-827-688-7 Sequence 7, Appli
21	4106	45.1	9030	1	US-08-324-977-13 Sequence 13, Appli
22	4106	45.1	9030	2	US-08-384-616-13 Sequence 13, Appli
23	4106	45.1	9030	2	US-08-904-686A-13 Sequence 13, Appli
24	4106	45.1	9030	3	US-09-315-850-13 Sequence 13, Appli
25	4106	45.1	9416	1	US-08-324-977-1 Sequence 1, Appli
26	4106	45.1	9416	2	US-08-384-616-1 Sequence 1, Appli
27	4106	45.1	9416	2	US-08-904-686A-1 Sequence 1, Appli

28	4106	45.1	9416	3	US-09-315-850-1 Sequence 1, Appli
29	4106	45.1	9416	3	US-08-823-895A-27 Sequence 27, Appli
30	4103.4	45.1	9401	2	US-08-432-693-1 Sequence 1, Appli
31	4103.4	45.1	9416	3	US-08-811-566-19 Sequence 19, Appli
32	4103.4	45.1	9416	3	US-09-034-756-19 Sequence 19, Appli
33	4101.4	45.1	9595	3	US-09-014-416-4 Sequence 4, Appli
34	4093.2	45.0	11076	4	US-09-539-601-1 Sequence 1, Appli
35	4093.6	44.9	8987	3	US-08-444-818-137 Sequence 137, App
36	4083.6	44.9	11076	4	US-09-539-601-25 Sequence 25, Appli
37	4082	44.9	11076	3	US-09-539-601-19 Sequence 19, Appli
38	4080	44.8	9416	3	US-08-823-895A-26 Sequence 26, Appli
39	4080	44.8	9416	4	US-10-104-966-13 Sequence 13, Appli
40	4077.2	44.8	11076	4	US-09-539-601-31 Sequence 31, Appli
41	4048.4	44.5	9472	4	US-08-150-204R-96 Sequence 96, Appli
42	4014.8	44.1	9185	3	US-08-444-818-122 Sequence 122, App
43	4014.8	44.1	9185	3	US-08-444-818-123 Sequence 123, App
44	3723.6	40.9	3970	1	US-07-925-695-3 Sequence 3, Appli
45	3678.4	40.4	8316	3	US-08-444-818-88 Sequence 88, Appli

ALIGNMENTS

RESULT 1
US-07-925-695-1
Sequence 1, Application US/07925695
Patent No. 5428145
GENERAL INFORMATION:
APPLICANT: OKAMOTO, Hiroaki
INVENTOR: NAKAMURA, Tetsuo
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOME,
POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
TITLE OF INVENTION: DETECTION SYSTEMS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Beveridge, DeGrandi, Wellacher & Young
STREET: 1850 M Street, N.W., Suite 800
CITY: Washington
STATE: D.C.
COUNTRY: US
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/925,695
FILING DATE: 19920807
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 287402/91
FILING DATE: 09-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 360441/91
FILING DATE: 05-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Wellacher, Robert G.
REGISTRATION NUMBER: 20,531
REFERENCE/DOCKET NUMBER: 06/87-48009
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2811
TELEFAX: (202) 659-1462
TELEX: WOI 64470
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9589 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-925-695-1
Query Match 93.4%; Score 8494.2; DB 1; Length 9589;

Best Local Similarity 75.3%; Pred. No. 0; Mismatches 6849; Conservative 1872; Mismatches 378; Indels 0; Gaps 0;	
Oy	1 ATGAGCAAAATCCCTAAACCTCAAGAAACCAAAAGAAACACCAACCGTCGGCCCAA 60
Db	341 AUGAGCAAAAUCCUAAACCCUAAAGAAAACCAAAAGAAACCAACCGTCGGCCCAA 400
Oy	61 GACGTAAGTTCCGGGCGGCGGCAGATCGTTGGCGGAGTATCTTTGGCCGCGAGG 120
Db	401 GACGTUAAAGUUTCCGGGCGGCGGCAGAUCCGUUGCGGAGUAAUCUUGUUGCCGCGAGG 460
Oy	121 GGGCCAGAGTTGGGTGTGTCGGCCCAAGAAAGACTTGGAGCGGTCCCAACCACTGGA 180
Db	461 GGCCCAAGUUGGGGUGCGCGCAAGAAAGACUUGGAGCGGUCGCAACCAAGGAG 520
Oy	181 AGGCGCCAGCCCATCTTAAAGATGGGCTCACTGGCAAACTCTGGGAGAAAACAAGA 240
Db	521 AGGCGCCAGCCCAUCCUAAAGAUCCGCGCUCACUGGCAAAUCCUGGAGAAAACAAGA 580
Oy	241 TACCCCTGGCCCTATACGGGAAAGAGGAGCTGGCTGGGCGAGATGGCTCTGTCGCC 300
Db	581 UACCCUGGCCCCCUAAACGGGAAUAGGGAUCCGGCTUGGCGAGAUUGGCUUCUCCGCC 640
Oy	301 CGAGGTTCCGCTCCCTCTTTGGGCCCCCAATGACCCCGGCAUAGGTGCGGCAAGTGGT 360
Db	641 CGAGGUTCCCGCCUUCUUGGGGCCCCCAUAGACCCCGGCAUAGGUCGCGCAAGUGGGU 700
Oy	361 AAGGCAATGATACCTCTAACGCGGGCTTTGCGCGCATGGGGGATACCTGTGCTG 420
Db	701 AAGGCAATGATACCTCTAACGCGGGCTTTGCGCGCATGGGGGATACCTGTGCTG 760
Oy	421 GGGCGCCCGCTCGCGGCGGTGCGCAAGACTCTCGGCAUAGGCTGAGAGTCTGAGAGAC 480
Db	761 GGGCGCCCGCTCGCGGCGGTGCGCAAGACTCTCGGCAUAGGCTGAGAGTCTGAGAGAC 820
Oy	481 GGGGTAATTTTTCGAAAGGGAATTACCCGGTGTCTCTTTTATTTTCTGCTGGCC 540
Db	821 GGGGTAATTTTTCGAAAGGGAATTACCCGGTGTCTCTTTTATTTTCTGCTGGCC 880
Oy	541 CTGCTGCTGTGATACCAACCCCGGTCTCGCTGCGAAAGTGAAGAAATCATGATACGGC 600
Db	881 CUGCGUUCUGCAUACCAACCCCGGUCUCCGCTGCGAAAGTGAAGAAATCATGATACGGC 940
Oy	601 TACATGTGATCAAGCACTGCAACCAATGACAGATTAATCTGGAGCTTCAGGCTGTCTC 660
Db	941 UACAUGUGAGCAACAGACUGCAACCAUAGUAGCAUUAACUGGCAACUCCAGGCUUGUC 1000
Oy	661 CTCACGTCCTCCCGGTGTCTCCGTGCGAAGAAAGTGGGAAATGCATCTCACTGCTGATA 720
Db	1001 CUCCAAGTCCCGGGUGGUCGCGCGGAGAAAGUGGGGAAUACUUCUGGUCUGGAGUA 1060
Oy	721 CCGGCTACCGAAATGTGGCCGTGAGCGGCGCGGCGGCTCAAGCAAGGCTTGCGGAGC 780
Db	1061 CCGGCTACCGAAATGTGGCCGTGAGCGGCGCGGCGGCTCAAGCAAGGCTTGCGGAGC 1120
Oy	781 CACATGCAATGTTGTGATGTCGCGCACGCTGTGCTGTGCTCTCACTGAGGGAGCTCTC 840
Db	1121 CACATGCAATGTTGTGATGTCGCGCACGCTGTGCTGTGCTCTCACTGAGGGAGCTCTC 1180
Oy	841 TGGCGTGGGGTGTGTCCTGCAAGCCCAATGTTTCAATTGTCTGCGCGAGACCACTGGTTT 900
Db	1181 TGGCGTGGGGTGTGTCCTGCAAGCCCAATGTTTCAATTGTCTGCGCGAGACCACTGGTTT 1240
Oy	901 GTCCAAAGACTGATTTGCTCATCTAACCTGGATCATCATGAGCAACGCGATGGCATGG 960
Db	1241 GUGCAAGACTGATTTGCTCATCTAACCTGGATCATCATGAGCAACGCGATGGCATGG 1300
Oy	961 GACATGATGATGAATGTGTGCGCCACGCGTCAACATGATCTTGTGCTGACGATGCTGTCTC 1020
Db	1301 GACAUGAUGAUGAACUGGUCGCGCCACGCGTCAACATGATCTTGTGCTGACGATGCTGTCTC 1360
Oy	1021 CCCGAGGTCAATTAATGATCATCTTAAGCGGGCTCAATTGGGGCTCATGTTTGGCTTGGCC 1080
Db	1361 CCCGAGGTCAATTAATGATCATCTTAAGCGGGCTCAATTGGGGCTCATGTTTGGCTTGGCC 1140
Oy	1081 TACTTCTATGAGAGGAGGAGTGGGCAAAAGTGTGTGTGATCTTCTGTTGGCGCGCGG 1140
Db	1421 UACUUCUUAUGAGAGGAGGAGTGGGCAAAAGTGTGTGTGATCTTCTGTTGGCGCGCGG 1480
Oy	1141 GTGAGCGCGGCAACCAATCTGTTGGGGGTTCTGCGCGAGACCAACCGGCGCCCTCAC 1200
Db	1481 GTGAGCGCGGCAACCAATCTGTTGGGGGTTCTGCGCGAGACCAACCGGCGCCCTCAC 1540
Oy	1201 AGCTTAATTTGACATGGGCCCCAGGAGAAATCCAGCTGTGTTAACCAATGGCAGCTGG 1260
Db	1541 GGCATGUGUCUCCUUGGUGCCAGGAGAAAUCACGCUACUACAACCAUUGGCAUUGG 1600
Oy	1261 CACATCAACGCGACCCGCCCTGAATCTCAATGATCTCTTGCACACCGGCTTTATGCGCTCT 1320
Db	1601 CACATCAACGCGACCCGCCCTGAATCTCAATGATCTCTTGCACACCGGCTTTATGCGCTCT 1660
Oy	1321 CTGTTTACACCCGACAGCTTCAACTGTGTGAGATGTCTCCGAACGCAATGTCGCTCGCC 1380
Db	1661 CUGUUCUACACCCGACAGCTTCAACTGTGTGAGATGTCTCCGAACGCAATGTCGCTCGCC 1720
Oy	1381 AGTATGAGGCTTCCGGGTGGGAGTGGGCGGCTTGCATTAATGAGGATTAATGACCAAT 1440
Db	1721 AGTATGAGGCTTCCGGGTGGGAGTGGGCGGCTTGCATTAATGAGGATTAATGACCAAT 1780
Oy	1441 CCAGAGATATGAGACCTTATTTGCTGAGCACTAACCAAGGAGTGTGGCGTGTCTCC 1500
Db	1781 CCAGAGATATGAGACCTTATTTGCTGAGCACTAACCAAGGAGTGTGGCGTGTCTCC 1840
Oy	1501 GCGAAGACTGTGTGTGGGCCCAAGTGTACTGTTTCAACCCCAAGCCCAATGATGTGGACG 1560
Db	1841 GCGAAGACTGTGTGTGGGCCCAAGTGTACTGTTTCAACCCCAAGCCCAATGATGTGGACG 1900
Oy	1561 ACCGACAGGCTTGGAGCGGCCCACTTAACGTGGGGGAGAAATGAGACAGATGTCTCTTA 1620
Db	1901 ACCGACAGGCTTGGAGCGGCCCACTTAACGTGGGGGAGAAATGAGACAGATGTCTCTTA 1960
Oy	1621 TTGAACGACACTGACCAACCGCTGGGGTCAATGGTTGCGGTGACATGATGAACTTCTTC 1680
Db	1961 TTGAACGACACTGACCAACCGCTGGGGTCAATGGTTGCGGTGACATGATGAACTTCTTC 2020
Oy	1681 GGTCAACCAAGACTTGGCGGCGCAACCTTGGCTGATCTAGACTGATCTTCAACCGCACG 1740
Db	2021 GGTCAACCAAGACTTGGCGGCGCAACCTTGGCTGATCTAGACTGATCTTCAACCGCACG 2080
Oy	1741 ACCGACCTGTTGGGCCCAAGGACTGTTTGAAGAGCACTGATCAACTTACCTTCAAA 1800
Db	2081 ACCGACCTGTTGGGCCCAAGGACTGTTTGAAGAGCACTGATCAACTTACCTTCAAA 2140
Oy	1801 TGGCGCTTGGGCGCTTGGCTCAAGCAAGTGTGCTGATGATCACTTACAGGCTCTGG 1860
Db	2141 TGGCGCTTGGGCGCTTGGCTCAAGCAAGTGTGCTGATGATCACTTACAGGCTCTGG 2200
Oy	1861 CATTAACCTTGCACAGTTAACTATACCATTTCAAAATTAAGATGTATGTGGAGGGGTT 1920
Db	2201 CATTAACCTTGCACAGTTAACTATACCATTTCAAAATTAAGATGTATGTGGAGGGGTT 2260
Oy	1921 GAGCAGAGGCTCAAGGCTGATGCAATTTCACTCGTGGGAGTGTGCAACTTGGAGGAC 1980
Db	2261 GAGCAGAGGCTCAAGGCTGATGCAATTTCACTCGTGGGAGTGTGCAACTTGGAGGAC 2320
Oy	1981 AGAGACAGAAAGTCACTGTCTCTTTTGTGCACTGCAACGGAATGGGCAATTTTACT 2040
Db	2321 AGAGACAGAAAGTCACTGTCTCTTTTGTGCACTGCAACGGAATGGGCAATTTTACT 2380
Oy	2041 TGTCTTATCTGGAAGCTGCGCTGCTGTGTGATGTGTCTTCACTTCCCAAAACATC 2100
Db	2381 TGTCTTATCTGGAAGCTGCGCTGCTGTGTGATGTGTCTTCACTTCCCAAAACATC 2440
Oy	2101 GTGAGGTAATTAATTAATGATGCTATCACTGAGCTCAAAATTAATGATGTCGATGG 2160
Db	2441 GTGAGGTAATTAATTAATGATGCTATCACTGAGCTCAAAATTAATGATGTCGATGG 2500

4661 ACCAAGCCGCUAGACGGGGUUAUCUGAGACUUGACUCGUGAUCGACUGCAAGUA 4720
4381 GGGGCACTCAAGTTGTAGACTTTAGACCCCACTTACCATTAACACACAGATT 4440
4721 GGGGACUCAAAGUUGUAGACUUCAGCUGGACCCACAUUCACAUACACACAGCU 4780
4441 GTCCCTCAAGACGCTGTCTCACTAGACGAGCGCGGGGTCCAGCGGTAGGGAGACTG 4500
4781 GUCCCUCAAAGCGCTGUCUACAGUAGCCAGCGCGGGGCGCACGGGGACGGGAGAGCUG 4840
4501 GGCATTTATAGGTATGTTTCCATGTGTAGAGGAGCGCTTAGAGAAATGTTTGAACGTGTAGT 4560
4841 GGUUUUUUAGUUAUUGUUUCCACUGUGUAGGAGCGCTUCAGAAAGUUUUGACUGUAGUG 4900
4561 CTGTGTAGGTCTACGACGACGAGGGGCGCATGTGTATGACTCACACCTCGGAGACACC 4620
4901 CUCUCGAGUGUACGAGACGAGGGGCGCAUGGUUAGUCUACACCGCGGAGACACC 4960
4621 GTCAGGCTCAGGGGCTATTTTCAACACGCGCGGTTTCTGTGTGCGCAAGACCATCTTAG 4680
4961 GUCAGGCTUCAGAGCAUUUUUCAAACACUGGUUUUCUGUGGCAAGACCAUUCUUGAG 5020
4681 TTTTGGGAGGCAGTTTTCACCGGCTCTACACATAGTGCCTTCTTCCCAACA 4740
5021 UUUUGGAGGCAUUUUUCCAGCGGCTUCACACAAUAGUAGCCCAUUCUUCCCAACA 5080
4741 AAGCAATCGGGGAAAAATTTGCGATCTTAACAGCTACGAGCTACAGTGTGCTAGG 4800
5081 AAGCAUUGGGGAAAAUUUUGCAUUAUUAAGCTUACAGGCUACAGUUGCGCUGAG 5140
4801 GCCAAAGCCCCCGCTCTGTGGAGCGTCACTGTGGAAGTGTGTACTGACTGAAGCCC 4860
5141 GCCAAAGCCCCCGCTCTGTGGAGCGTCACTGTGGAAGTGTGTACTGACTGAAGCCC 5200
4861 ACACCTGTGTGGGCCCCCAACACTCTCTGTACCGCTTGGGCTCTGTGTAACAAAGAGTCAAC 4920
5201 ACACCTGTGTGGGCCCCCAACACTCTCTGTACCGCTTGGGCTCTGTGTAACAAAGAGTCAAC 5260
4921 CTCACACATCCCGTATCGCAAAATATCATGCGACCTGATGCAAGCGGACCTTGAAGTCAATG 4980
5261 CUCACGACUCCUGUACCAAAUUAUACUCCGACUCUACAGGCAAGCCGACCUUAGAGUACUG 5320
4981 ACCAGCAGATGGGTCTTGGCAGAGGGAGTCTTGGCGGCGCTGCGCGCTATTGCTGTGCG 5040
5321 ACCAGCAGUUGGUUUCUUAUGGUGGGGUGUUGGCGGCGGCTGUGCGGCUACUGCGCG 5380
5041 ACCGGGTGTGTTGATCATGCGGCGCTTGCATTAACACAGCGAGCGCTGTGTCGCG 5100
5381 ACCGGGUUUGUUGAUCAUCGCGCUGUACAGUUAACAGCGAGCGGUGUGCACCG 5440
5101 GACAAGAGAGTCTCTATAGAGCTTTTGTATGATGAGAGAAATGCTCTTAGAGGGGCT 5160
5441 GACAAGAGAGUCCUUAUGAGGCUUUUAUGAUGAGUAGAAUUGGCUUAGAGGCGCU 5500
5161 CTCATTGAAAGAGGGGACGCGATAGCCGAGATGTAAGTCCAGATCCAAAGGCTTATTG 5220
5501 CUCAUUGAAAGGGGACGCGGUAAGCCGAGUAGCUUAGUCCAAAGUCCAAAGGCUUATUG 5560
5221 CAGCAAGCTTCCAAACAAGCTCAAGACATTAACCCACTGTGCAAGGCTTCAATGCGCAAG 5280
5561 CAGCAAGCTTCCAAACAAGCTCAAGACATTAACCCACTGTGCAAGGCTTCAATGCGCAAG 5620
5281 GTAGAACTTTCTGGGCCCCAAACATGTGAACTTCAATTAAGGCGATCCAAATCTGCA 5340
5621 GUAGAGCAUUCUGGGCCAAACACAUUGGAAUUCUACACGCGGCAUUAUACUUGCA 5680
5341 GGACTATTAACACTGCGCAGAGGAACTCTGCAAGTACTTCATGATGAGCTTCAAGTGCAGC 5400
5681 GGAUUAUCAAACUCCGACGAGGAACTCUGUAGUUCUACAUUGGCAUUAUAGUCCGCGC 5740
5401 CTCACCAATCCGCTGTCAACAGCACTATCTTCTCAACATTTTGGGGGCGTGTGCTA 5460

5741 CUCACAGUCUGUUGUCAAUAGACCAUUAUUCUUUCAAUAUUUUUGGGGCGUGCUA 5800
5461 GCATCCCAAAATTTGACACACCGCGGGGCGCACTGCTGTGTTGATGAGTCAATGAGGA 5520
5801 GCAUCCCAAAUUBCGCTCCGCGGAGGCUACCGGCUUUGUGUUCAGUUGGCGUGUGGG 5860
5521 GTCGCGTAGGCAAGTATAGGCTTAAAGTGTCTAGTGAACATCTTGCAGAGGTATGTT 5580
5861 GCGCGGUAAGGACAGAUAGGCUUGGUUAGUUGUCUGGGAUUCUGGACAGGUAUUGU 5920
5581 GCGGCAATTTCCGGGGGCTCTCGTCCGATTCAAGATATGTCTTGGGAGAAAGCCCTCCAG 5640
5921 GCGGCAUUCGCGGCGCUUCGUGCGCAUUAUUAUUAUUGUCGAGAAAGCCUCCUAG 5980
5641 GAGAGTGTGCTCAACTTGTGCTGCTGGAATTTGTCTCCGGGATGCTTGTGATGAGGAGTC 5700
5981 GAGAGUUGUUCACACUUGCUGCGGGAUUAUCUUCUCCCGGUGCCUGUGUGGAGUUC 6040
5701 ATCTGCGGCGCAATTTGCGCGCAACAGTGGGACCGGGGAAAGCGCCCTCCATGAGATG 5760
6041 AUCUGCGGCGCAUUCUGCGCGACAGUGGGAACCGGGGAAAGCGCUCUCCAAUUGAGUG 6100
5761 AATAAGCTAATTTGCTTGTCCAGAGGAATCAAGTGGCCCCCACTCACTGAGAGC 5820
6101 AAUAGGCUCAUUBCGCTCUCUUGCUUCAGAGGAACACAGUGCGCCCACTCAUAGUGAGC 6160
5821 GAGTCGAGTGCCTGCGAGCGGTGACCCAACTACTTGGCTCTCTTACATAACACAGCTG 5880
6161 GAUUCGAGUUCGUGUGAGGUGUAGCCAAUUAUUGUGUCCUUAUUAUACCAACCTG 6220
5881 CTCAGAAAGTCTCAACACTGAGATTACTGAGACTGCCCCATGCCAGCGGCTGTG 5940
6221 CUCAGAGACUCCACAACTGAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 6280
5941 CTCGCGAGTGTGGGACTGGGTTGCAACCATCTTAACAGACTTAAATTTGAGTGAAC 6000
6281 CUCGCGAGUUGGAGAUUGGAGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 6340
6001 TCCAAATTTATCCCAAAAGATGCGCGGCTCCCTTGTCTCTGTGCAAAAAGGGTACAAG 6060
6341 UCCAAUUGUUCGCAAAAGUAGUGCUGUGUUCUCCUUUAUUCUUGUCAAAGGGGUAACAAG 6400
6061 GCGGTGTGGGCGGCACTGCGATCATGACCAACGCTGTCTTGGCGGCGCAATATCTCT 6120
6401 GCGGUGGCGCUGGCACTGAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 6460
6121 GCGAAGTCCGCTTGGGCTCTAGAGATTAAGAGGCGCTTAAGACCTGATGAATATCTG 6480
6461 GGCAGUUCGCGCUGGCUUCAGAGAAUUAAGGCGCCAAACCUAGAUUAUUCUG 6520
6181 CAGGAGCACTTCTCTATCAATTTGTTACAGGAGGCGCAGTGCCTGCGCAAAACCGCGCGCA 6240
6521 CAGGAGCACTTCTCTATCAATTTGTTACAGGAGGCGCAGTGCCTGCGCAAAACCGCGCGCA 6580
6241 AACTTTAAGTGGCCTATGAGAGGTGGCGGCTCAAGATGACCGGAGGTGAGCGAGC 6300
6581 AACUUUAAGUCCCAUUCUGAGAGGCGGCGCCUACAGAGUACCGGAGGAGUAGCGAGCAC 6640
6301 GGGATCAACCACTAACAATTAACAGGACTCACCATGATTAATTTGAAAGTCCCTGCCAATCA 6360
6641 GGGUCAAACCAUUAACAGAGACUUAACCAUAGUUAUUAUUAUUAUUAUUAUUAUUAUUA 6700
6361 CCTCTTCCGAGTTCTTTCTCTGAGTGAACGAGAGTCAAGTCAATGAGTTTGCCTCCACA 6420
6701 CUUUUCACAAGUUCUUUUCUUGGUGAGAGGAGUCCAGUUCUUAUUGUUGCCCCCAUA 6760
6421 CCGAAGCGTTTTCGCGGATGAGGTCTGCTTCTGCGTTTGAATTTATTTATTTGCTGTC 6480
6761 CCGAAGCGUUUUUUGGAGUAGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 6820
6481 GGGTCCAGGCTTCTGCGAGCCTGGAACCGGACACAGAGCTATTGATGTCATGCTTACA 6540
6821 GGGUCCAGUCCUUGCGAUUCUAGACUUAACAGAGUUAUUAUUAUUAUUAUUAUUAUUA 6880

OY	6541	GATCATCTTCAATATCAAGCGGAGACATGACACGCGCGCTTTAGCGCGGGGTCAACCCCA	6600
Db	6881	GACCCAUCCCAUUAUACAGCGAGACUACACGCGCGCTUUUGGACACGGGGGUCACCCCG	6940
OY	6601	TCCGAGCGAACTCCTCGCGAGCGACGATATCGGACATGACATGCTGCGAGCCACATGCAAC	6660
Db	6941	UCCGAGCGAACTCCTCGCGAGCGACGATATCGGACATGACATGCTGCGAGCCACATGCAAC	7000
OY	6661	ACCCACGCGAAAAGCTTATGATGTGGA CATGTGGATGCTTAACTGTTTCATGGGGCGAT	6720
Db	7001	ACCCACGCGAAAAGCTTATGATGTGGA CATGTGGATGCTTAACTGTTTCATGGGGCGAT	7060
OY	6721	GTTGACTCGGATAGAGTCTGGGGTCCAAAATGGTGTCTTGGA CTCTCTGACCAATAAGTCTC	6780
Db	7061	GUGACCCGGAUAGAGUCTUGAGUCCAAAAGGUGUGUUGCGACUUCUUGACCTUGACCAAUUGUC	7120
OY	6781	GAAGAAAGAGACGACCTTGAAGCTTTCGATACCATACAGATACATGCTTCCCAAGAGAGS	6840
Db	7121	GAAGAAAGAGACGACCTTGAAGCTTTCGATACCATACAGATACATGCTTCCCAAGAGAGS	7180
OY	6841	TTCCCAACCAAGCTTATACGCGCTCGGGGCATGCGCTGATTACAA CCAACGCGCTTGTGAATCG	6900
Db	7181	UUCCCAACCAAGCTTATACGCGCTCGGGGCATGCGCTGATTACAA CCAACGCGCTTGTGAATCG	7240
OY	6901	TGGAAGAAAGCGCAGATTACCAACGCGGCACCTGTTGGGGCTGTGCTTCCCTCTCTAGAG	6960
Db	7241	UGGAAGAAAGCGCAGATTACCAACGCGGCACCTGTTGGGGCTGTGCTTCCCTCTCTAGAG	7300
OY	6961	AAAAACCCGACGCGCTCCCTCCCAAGAGAGCGCGGACATGAGGCTTAA GTAGAGATCCATA	7020
Db	7301	AAAAACCCGACGCGCTCCCTCCCAAGAGAGCGCGGACATGAGGCTTAA GTAGAGATCCATA	7360
OY	7021	GGAGATGCGCTTCAACAGCTGGCCATTAAAGTCTTTGGCCAGCGCCCTCCCAAGGGGGAT	7080
Db	7361	GCAAGUGCCCAACACAGCTGGCCATTAAAGTCTTTGGCCAGCGCCCTCCCAAGGGGGAT	7420
OY	7081	TCAGAGCTTTTCAACGCGGGGCGGGGCGCTCCGATTTCCGGCACTGACAGCGCTCTGATAGS	7140
Db	7421	UCAGAGCTTTTCAACGCGGGGCGGGGCGCTCCGATTTCCGGCACTGACAGCGCTCTGATAGS	7480
OY	7141	TTGGCCCTTTTGGGAGACAGGTTCCATCTCTTTCATGCCCCCTCGAGGGGAGACTTGGG	7200
Db	7481	UUGGCCCTTTTGGGAGACAGGTTCCATCTCTTTCATGCCCCCTCGAGGGGAGACTTGGG	7540
OY	7201	GATCACAAGCTGAGAGCTGAGCAGGTAGAGGCCCAACCCCCCTCCAGGGGGGGGTGGCA	7260
Db	7541	GAUCCACAAGCTGAGAGCTGAGCAGGTAGAGGCCCAACCCCCCTCCAGGGGGGGGTGGUA	7600
OY	7261	GCTCCGAGCTCGACCTCGGGGTCTGTGTTACTTGTCTCCGAGAGACGATCCGTGCTGTG	7320
Db	7601	ACCCCGAGCTCGACCTCGGGGTCTGTGTTACTTGTCTCCGAGAGACGATCCGTGCTGTG	7660
OY	7321	TGCTGCTCCATGTCATCTCCTGGAACCGGGGCTTATTAATCCTTGTATGTCGGAAGAG	7380
Db	7661	UGCTGCTCCATGTCATCTCCTGGAACCGGGGCTTATTAATCCTTGTATGTCGGAAGAG	7720
OY	7381	GAGAAAGTTTACCGATTAAACCTCTTAGCAACTCCCTGTTGCGATATCA CAACAAGSTGTAC	7440
Db	7721	GAATAAGTTTACCGATTAAACCTCTTAGCAACTCCCTGTTGCGATATCA CAACAAGSTGTAC	7780
OY	7441	TGTTACCAACAAAGAGCGCTCTCACTAAAGGCTTAAAAAGTTAACTTTTGATAGATGCAA	7500
Db	7781	UGTTACCAACAAAGAGCGCTCTCACTAAAGGCTTAAAAAGTTAACTTTTGATAGATGCAA	7840
OY	7501	GTGCTCGACTCTCTACGACTCACTGACTTTTAAAGAGACATTAACTAGTCAGGCGCTCAAGSTC	7560
Db	7841	GCCTCGACGCTCTCTACGACTCACTGACTTTTAAAGAGACATTAACTAGTCAGGCGCTCAAGSTC	7900
OY	7561	ACCGCAAGGCTCTCAACCATGAGAGAGGCTTCCAGTTAAACCCCAACCCCATTTCTGCAAGA	7620
Db	7901	ACCGCAAGGCTCTCAACCATGAGAGAGGCTTCCAGTTAAACCCCAACCCCATTTCTGCAAGA	7960

QY	7621	TCCTAAATATNGGGGTTTGGGGCTTAAGAGAGGTCCGACGTTTGTCCGGAGAGGGCCGTTAACAC	7680
Db	7961	UCCAGUAAUGGGUUBGGGCUAAAGAGAGGUCGCGAGUUUUUCCGAGAGCCGUAAC	8020
QY	7681	ATCAAAGTCCGCTGTGGAGAGACCTCCGAGAGAGACTCAGAAACACAAATTCCTCAACACATT	7740
Db	8021	AUCAAAGUCCGUGUUGGAGAGGACCUCCUGGAAAGACACAAACACCAUUCUACAAACAU	8080
QY	7741	ATGGCCAAAATAGAGGTGTTCTGCGTGAACCCCAACCAAGGGGGGCAAGAAAGCAGTCCG	7800
Db	8081	AUGGCCAAAAAUGAGGUGUUGCGUGGAGACCCCAACCAAGGGGGGUAAGAAAGACGUCGC	8140
QY	7801	CTTAATCGTTTATACCTTGACCTTCGGCGGTCAAGGGTCTGGGAAAGATATGGCCCTTTATGACATT	7860
Db	8141	CUUACCGUUBACCUCGACCUCCGCGCGUACAAGGUCUGGCAAAAAUUGGCCCUUUUUAUUAU	8200
QY	7861	ACACAAAAAATTCTCTCAGGCGGTGATGAGGGGGGCTTCTTATGATTTCCAGTATTCACCGCT	7920
Db	8201	ACACAAAACCTUCCUCAGCGGUGAUGGGGGCUUUCUUAUGAUUCCAGUACUCCCCGCU	8260
QY	7921	CAGCGGGTATAGATTTCTCTTGAAGCATGGGCGGAAAGAAAGAACCCCTATGGGTTTTTCG	7980
Db	8261	CAGCGGGUGAGUUYUCUUGAAGGCAUGGGCGGAAAGAAAGAACCCUAUUGGUUUUUG	8320
QY	7981	TATGATATCCGATGCTTTGACTCAACCGTCACTGAGAGACATCAGACTGAGAGTCC	8040
Db	8321	UAGUAGAACCCGAGUGCUUUAUCUAAACCGUACAUGAGAGAGACAUCCAGCATUGGAGAGUCC	8380
QY	8041	ATATATCCGGGCGCTGCTCTCTTGCGCGAGAGAGGCCACATCGGCATACCTCGCTAATAG	8100
Db	8381	AUAUUAUCGGGCUUUGUCCUUGCCGAGAGAGGCCACACUCCGACUAAUCACUCAGCAGUAG	8440
QY	8101	AGACTTATCGTGGAGAGGGCCTTATGTTTCAACAGCAAGGGGCCAAAATTGCGGGTTCAGGCGT	8160
Db	8441	AGAUCUUAACGUGGAGAGGGCCCAUGUUCAAACAGACAGGCGCAACUCUGCGGUAUACGGCGU	8500
QY	8161	TGCGCGCCAGCGGGGTGCTCAACAATAAGCATGGGGAAACCATCATCATGCTTATGTTGAAA	8220
Db	8501	UGCCCGCGCACGAGGGUGCUUACCUAGCAGGGGAAACACAUCAUAGCUUUAUGUAAA	8560
QY	8221	GCGTTAAGGGGCTTGTAAAGCTGCAAGGATTAATCGGCCCCCAATGCTGTATGCGGCGAT	8280
Db	8561	GCUUAGGGGCGCUUAGACUGGCAAGAAAUUUGCGCCCAUAUGUGUAUUGCGCGAU	8620
QY	8281	GACTTGTGTTTCATCTCAGAAAGCCAGAGGAGCCAGAGAGAGAGCGGCGAACTGAGAGCC	8340
Db	8621	GACUUGUGUUCUAUCUCAGAGAGCCAGAGGAGCCAGAGAGAGAGAGAGCGGAAACUUGAGACC	8680
QY	8341	TTACAGAGAGCTTATGACGAGGTATTCTGCGCCCTCTGTGTGACCCCCCGAGCCGAGAT	8400
Db	8681	UUCAAGAGGGCUAUGACCAAGUAUUCUGCCCUUCCUGUGAGACCCCCCGAGCCGAAUUV	8740
QY	8401	GATCTGAGCTGATTAACATCTTGCTTCCTCAAAATGTGTCTGTGGCGCTGGGCCCAAGAGC	8460
Db	8741	GACCTGGAGCUGAUUACAUCUUGCUCUCAAAUUGUUGUCUGUGGGGCCACAAAGC	8800
QY	8461	CGCGCGAATCTACTGACCAAGAACCTTATCCATCCAAATGGCCCGGGCTGCTGGGAA	8520
Db	8801	CGCGCGAAUACUACCUAGACAGAGAACCCUACACUCCCAUUGCGCCGGGCGUGCUUGGAA	8860
QY	8521	ACAGTTTAAACACTCCCTGTCATAATTTCATGAGCTGGGAAACATCATCCAGTACGCCCGAGC	8580
Db	8861	ACAGUUAACACUCCCGUGUACAUAUUGCUAAGGAAACAUUCCAGUAACGCCCCAAAC	8920
QY	8581	ATATGGGCTCCGATGCTCTGTGATGACACATTTCTTCTCATTTCTCATGCTCAAGACACG	8640
Db	8921	AUAUUGGCGUCGUAUGUCCUGAUGACACACUUCUUCUACUUGCCCAAGUAUACU	8980
QY	8641	CTGAGACGAAACCTCAACTTTGAGATGTAACGAGAGGGGTACTCTCCGTGATGCCCTTGAGC	8700
Db	8981	CUGAGACGAAACCUUACUUAUGAUGUAGAGGAGCGGUGUACUCCGUGAGUCCCTUUGAGC	9040
QY	8701	CTCCACGCTATATTTGAAGGTTACATGAGGCTTGAAGCTTTTCTCTGCAACATACACT	8760

Db 9041 CUCCACGCAUAAUUAAGGUAACAGCGGCUUACCGCUUUCUCUCGACACAUACU 9100
Qy 8761 CCCCAGCAATGACACGCGGTGGCTTTCAGCCCTCAGAAAATTGGGGGCGCCACCCCTCAGA 8820
Db 9101 CCCCAGCAATGACACGCGGTGGCTTTCAGCCCTCAGAAAATTGGGGGCGCCACCCCTCAGA 9160
Qy 8821 GCGTGAAGAGCCGCGGACGTCGACGACGAGGCGCTCCCTCATCTTCCGTGGGGGAGAGCG 8880
Db 9161 GCGTGAAGAGCCGCGGACGTCGACGAGGCGCTCCCTCATCTTCCGTGGGGGAGAGCG 9220
Qy 8881 GCGTGAAGAGCCGCGGACGTCGACGAGGCGCTCCCTCATCTTCCGTGGGGGAGAGCG 8940
Db 9221 GCGTGAAGAGCCGCGGACGTCGACGAGGCGCTCCCTCATCTTCCGTGGGGGAGAGCG 9280
Qy 8941 TTGCGGAGAGCGCGCTCTCTGATTTATTCAGCTGTTCACCCGTGGGCGCGGCGGCG 9000
Db 9281 UUGCGGAGAGCGCGCTCTCTGATTTATTCAGCTGTTCACCCGTGGGCGCGGCGGCG 9340
Qy 9001 GACATTTATACAGCGGTGTCGCGTCCCGACCCCGCTTATTCCTTTGGCTTACTCTTA 9060
Db 9341 GACATTTATACAGCGGTGTCGCGTCCCGACCCCGCTTATTCCTTTGGCTTACTCTTA 9400
Qy 9061 CTTTGTAGGCGGTGAGCGCTTTTCTTACTCTCCCGCTGG 9099
Db 9401 CUUUUUGAGGGGUAGCGCUUUCUACUCCCGCTCGG 9439

RESULT 2
US-07-925-695-2
Sequence 2, Application US/07925695
Patent No. 5428145
GENERAL INFORMATION:
APPLICANT: OKAMOTO, Hiroaki
APPLICANT: NAKAMURA, Tetsuo
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOME.
TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: Beveridge, Degrandi, Wellacher & Young
STREET: 1850 M Street, N.W., Suite 800
CITY: Washington
STATE: D.C.
COUNTRY: US
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/925,695
FILING DATE: 19920807
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 287402/91
FILING DATE: 09-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 360441/91
FILING DATE: 05-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Wellacher, Robert G.
REGISTRATION NUMBER: 20,531
REFERENCE/DOCKET NUMBER: 06/87-48009
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2811
TELEFAX: (202) 659-1462
TELEX: WUI 64470
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 9589 base pairs
TYPE: NUCLEIC ACID

STRANDEDNESS: single
TOPOLOGY: linear
US-07-925-695-2
Query Match 93.4%; Score 8494.2; DB 1; Length 9589;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 8721; Conservative 0; Mismatches 378; Indels 0; Gaps 0;
Qy 1 ATGAGCAAAATCTTAAACCTTAAGAAAAACCAAAAGAACCAACCGTGGCCACAA 60
Db 341 ATGAGCAAAATCTTAAACCTTAAGAAAAACCAAAAGAACCAACCGTGGCCACAA 400
Qy 61 GACGTTAAGTTTCCGGGCGCGCCGAGATCTGTTGGGAGTATCTTGTGGCGGACAG 120
Db 401 GACGTTAAGTTTCCGGGCGCGCCGAGATCTGTTGGGAGTATCTTGTGGCGGACAG 460
Qy 121 GGGCCAGGTTGGGTGTGGCGCGACCAAGAAAGACTTGGAGCGGTCCAGGCAAGTGA 180
Db 461 GGGCCAGGTTGGGTGTGGCGCGACCAAGAAAGACTTGGAGCGGTCCAGGCAAGTGA 520
Qy 181 AGGCGCAAGCCATCCCTTAAAGATCGGCGCTCCTGAGCAATCTTGGGAAAAACAGGA 240
Db 521 AGGCGCAAGCCATCCCTTAAAGATCGGCGCTCCTGAGCAATCTTGGGAAAAACAGGA 580
Qy 241 TACCCCTGGCCCTTATAGGGAATAGAGGACTTCGGTGGGAGATGCTCTGTCCCGC 300
Db 581 TACCCCTGGCCCTTATAGGGAATAGAGGACTTCGGTGGGAGATGCTCTGTCCCGC 640
Qy 301 CGAGTTCCGTCCTCTTGGGGCCCAATGACCCCGGCAATAGTGGCGCAAGCTGAGT 360
Db 641 CGAGTTCCGTCCTCTTGGGGCCCAATGACCCCGGCAATAGTGGCGCAAGCTGAGT 700
Qy 361 AAGGTATGATACCTTAACGTGGGCTTTGCGGACCTCATAGGCTGCTGCTGCTG 420
Db 701 AAGGTATGATACCTTAACGTGGGCTTTGCGGACCTCATAGGCTGCTGCTGCTGCTG 760
Qy 421 GGGGCCCCGCTCGGCGGCGCTCGCAAGGCTTCGCGCAATGAGGCTGAGTACTCTGAGAG 480
Db 761 GGGGCCCCGCTCGGCGGCGCTCGCAAGGCTTCGCGCAATGAGGCTGAGTACTCTGAGAG 820
Qy 481 GGGGTTAATTTTGAACAGAGAACTTACCGGCTGCTCTTCTTCTTCTTCTTCTGAGG 540
Db 821 GGGGTTAATTTTGAACAGAGAACTTACCGGCTGCTCTTCTTCTTCTTCTTCTGAGG 880
Qy 541 CTGCTGTCTGATCAACACCCCGGTCTCGCTGCCGAAGTGAAGAACTACGACCGG 600
Db 881 CTGCTGTCTGATCAACACCCCGGTCTCGCTGCCGAAGTGAAGAACTACGACCGG 940
Qy 601 TACATGTGATTAAGCATGACCAATGACAGATTAACCTGGAGCTCCAGGCTGCTGTC 660
Db 941 TACATGTGATTAAGCATGACCAATGACAGATTAACCTGGAGCTCCAGGCTGCTGTC 1000
Qy 661 CTCACAGTCCCGGGGTGCTCCCGTGGAGAAAGTGGAGAAATGATCTCAGTGTGGATA 720
Db 1001 CTCACAGTCCCGGGGTGCTCCCGTGGAGAAAGTGGAGAAATGATCTCAGTGTGGATA 1060
Qy 721 CCGGTCTACCGAATGAGCGGTGACGCGGCGCCGCTTACGACGAGGCTTTGGAGG 780
Db 1061 CCGGTCTACCGAATGAGCGGTGACGCGGCGCCGCTTACGACGAGGCTTTGGAGG 1120
Qy 781 CACATGACATGTTTGTATGTCGCCGACGCTCTGCTGCGCTTACGAGGGGAGCTTC 840
Db 1121 CACATGACATGTTTGTATGTCGCCGACGCTCTGCTGCGCTTACGAGGGGAGCTTC 1180
Qy 841 TGGGTGGGGTGTATGCTCCAGCCAAATGTTTATTTCTCGCGGAGACCACTGGTTT 900
Db 1181 TGGGTGGGGTGTATGCTCCAGCCAGATGTTTATTTCTCGGCAACGACCACTGGTTT 1240
Qy 901 GTCCAAAGATGCAATTTCTTCATCTTACCTTGGTACATGACACCGCATGGCATGG 960
Db 1241 GTCCAAAGATGCAATTTCTTCATCTTACCTTGGTACATGACACCGCATGGCATGG 1300
Qy 961 GACATGATGATGAACGTGTGCGCCAGCGCTACCATGATCTTGGCGTATGCGGTGTC 1020

1301 GACATGATGATGAACTGCTGCGCCACGCGCTACCATGATCCCTGGGAGTACGCGATCGCGCTC 1360
1021 CCCGAGGTCATTATGACATCATTTAGCGGGGCTCATTTGGGGGCTGATGTTGCGCTTGACC 1080
1361 CCCGAGGTCATCATGACATCATTTAGCGGGGCTCATTTGGGGGCTGATGTTGCGCTTGACC 1420
1081 TACTTCTTATGACGAGGAGCGTGGGCGAAAGTCGTGTCATCTTCTGTTGGCGCGCGG 1140
1421 TACTTCTTATGACGAGGAGCGTGGGCGAAAGTCGTGTCATCTTCTGTTGGCGCGCGG 1480
1141 GTGAGCGCGCGCACCATTAATGTTGGGGTTCGCGCGGACGACCGCGCGCTCAAC 1200
1481 GTGAGCGCGGAAACCATTAACGTTGGGGTTCACCGGACATTAACGACGAGACCTCAAC 1540
1201 AGCTTATTTGACATGGGCGCCGACGACGAAATCCAGCTCTTTAAACCATGACGCTGG 1260
1541 GGCATGTTCTCCCTTGGTGGCGAGGACGAAATCCAGCTCATCAACCATGAGCATGGG 1600
1261 CACATCAACGCGCACCGGCTGGAATGCAATGCACTCTTGGACACCGGCTTTATGGCGCT 1320
1601 CACATCAACGCGCACCGGCTGGAATGCAATGCACTCTTGGACACCGGCTTTCTGCGCTCA 1660
1321 CTGTTCTACACCCACAGCTTCACTCGTCAAGATGTCGCGAACGATGTCGCGCGCGC 1380
1661 CTGTTCTACACCCACAGCTTCACTCGTCAAGATGTCGCGAACGATGTCGCGCGCGC 1720
1381 AGTATCGAGGCTTCCGGGTGGAGTGGGGCGCTTGGCAATGAGATTAATGTACCAAT 1440
1721 AGTATCGAGGCTTCCGGGTGGAGTGGGGCGCTTGGCAATGAGATTAATGTACCAAT 1780
1441 CCAGAGGATATGAGACCTTATGCTGGACATACCACAGGCGAGTGGGCTGCTCC 1500
1781 CCAGAGGATATGAGACCTTATGCTGGACATACCACAGGCGAGTGGGCTGCTCC 1840
1501 GCGAAGACTGTGTGGCCGACAGTGTACTGTTTCAACCCCGACGCGCATGTAAGTGGGACG 1560
1841 GCGAAGCTGTGTGGCCGACAGTGTACTGTTTCAACCCCGACGCGCATGTAAGTGGGACG 1900
1561 ACCGACAGGCTTGGAGCGCCCATTTACACGTTGGGGGAGAAATGAGACAGATGTTCTCTA 1620
1901 ACCGATATGACTTGGAGCGCCCATTTACACGTTGGGGGAGAAATGAGACAGATGTTCTCTA 1960
1621 TTGAAACGACTGACACACCGCTGGGGTGCATGTTCCGGCTGACAGTGAATGAATCTTCT 1680
1961 TTGAAACGACTGACACACCGCGAGGGATCATGTTCCGCTGACAGTGAATGAATCTTCTA 2020
1681 GGCCTACCAAGACTTGGCGGACACACCTGCGTACTAGAGCTGACTTCAACGCGACG 1740
2021 GGCCTACCAAGACTTGGCGGACACACCTGCGGATTAAGAGCTGACTTCAACGCGACG 2080
1741 ACCGACCTGTTGTGCCCCACGAGACTGTTTGAAGAGATCTGTATACCACTTACCTCAAA 1800
2081 ATGGAATGTTGTGCCCCACGAGACTGTTTGAAGAGATCTGTATACCACTTACCTCAAA 2140
1801 TGGGGCTCTGGGGCTCTGAGCTGACGCGCAAGGAGCTGATGCACTAACCTCAAGGCTGG 1860
2141 TGGGGCTCTGGGGCTCTGAGCTGACGCGCAAGGAGCTGATGCACTAACCTCAAGGCTGG 2200
1861 CATTACCCCTGACACAGTTAACTATACCATCTTCAAAATTAAGATGTAATGTTGGAGGGTT 1920
2201 CATTACCCCTGACACAGTTAACTATACCATCTTCAAAATTAAGATGTAATGTTGGAGGGTT 2260
1921 GAGCACAGGCTCACGGCTGTCATGCAATTTCACTGTGGGATCTTTCAACTTGGAGAGAC 1980
2261 GAGCACAGGCTCACGGCTGTCATGCAATTTCACTGTGGGATCTTTCAACTTGGAGAGAC 2320
1981 AGAGACAGAGTCAAGCTGTCTTGTGTGACCTCAACGGAATGGGCACTTTATCT 2040
2321 AGAGACAGAGTCAAGCTGTCTTGTGTGACCTCAACGGAATGGGCACTTTATCT 2380
2041 TGCTTACTCTGACCTGCGCGCTTGTGACCTGCTTCTCAACCTTCAACCAAAATCT 2100

2381 TGCACTTACTCGGACCTGCCGCGCTTGTGACATGCTCTTCTCCACCTCAACAAATCATC 2440
2101 GTGAGCGTACAAATTCATGATGAGCTTATGACCTGCGCTCAACAAATTCATGCTCCATGG 2160
2441 GTGAGCGTGAATTCATGATGAGCTTATGACCTGCGCTCAACAAATTCATGCTCCATGG 2500
2161 GAGTGGGTAAATCTTATTCCTGCTCTTGTAGCGGACGCGAGGGTTTCCGCTGCTTATGG 2220
2501 GAGTGGGTAAATCTTATTCCTGCTCTTGTAGCGGACGCGAGGGTTTCCGCTGCTTATGG 2560
2221 ATGCTATCTTGTGTGGCGACGCGCGGACGACATGAGAGAGCTGTATCTTGGACGCT 2280
2561 ATGCTATCTTGTGTGGCGACGCGCGGACGACATGAGAGAGCTGTATCTTGGACGCT 2620
2281 GCGAGCGCACTAGCTGACATGAGCTTCCATATTTTGTATCTTTTGTGCTGCTGG 2340
2621 GCGAGCGCACTAGCTGACATGAGCTTCCATATTTTGTATCTTTTGTGCTGCTGG 2680
2341 TACATCAAGGAGTGGGATGCGCCCTTACCTACCTATCCCTCACTGAGCTGTGGTCTTT 2400
2681 TACATCAAGGAGTGGGATGCGCCCTTACCTACCTATCCCTCACTGAGCTGTGGTCTTT 2740
2401 AGCCTACTGCTCTAGCATTTGCCCAACAGGCTTATGATGACGATCTGTGACATGCGC 2460
2741 GGCCTACTGCTCTAGCATTTGCCCAACAGGCTTATGATGACGATCTGTGACATGCGC 2800
2461 CAGATAGAGAGCGCTGCTGTGTGATGATCATCTCTTTACTCTACCCCGGGGTATAG 2520
2801 CAGATAGAGAGCGCTGCTGTGTGATGATCATCTCTTTACTCTACCCCGGGGTATAG 2860
2521 ACCCTTCTCAGCGGTTTTTGTGTGGTGTGCTATCTTGTGACCCCTGAGGGAAAGCTATG 2580
2861 ACCCTTCTCAGCGGTTTTTGTGTGGTGTGCTATCTTGTGACCCCTGAGGGAAAGCTATG 2920
2581 GTCGAGAGTGGGACCAACCTATGACAGTGGCGGCGGTGGATGAGATCATATGAGGCC 2640
2921 GTCGAGAGTGGGACCAACCTATGACAGTGGCGGCGGTGGATGAGATCATATGAGGCC 2980
2641 GTGGCCATATTCATACCAGTGTGTGTTGACATTAACCAAGTGGCTTGGGGGTCTT 2700
2981 GTGGCCATATTCATACCAGTGTGTGTTGACATTAACCAAGTGGCTTGGGGGTCTT 3040
2701 GGGCGTGTCTTACCTCTTAAAGGTTGCTTGAACGCGGTGCGTACTTGTGACAGGCTCAC 2760
3041 GGGCGTGTCTTACCTCTTAAAGGTTGCTTGAACGCGGTGCGTACTTGTGACAGGCTCAC 3100
2761 GCTCTACTAGAGATGTCACCATGACAGGATCTGCGGGGGGAGGTAAGTCAAGATG 2820
3101 GCTCTACTAGAGATGTCACCATGACAGGATCTGCGGGGGGAGGTAAGTCAAGATG 3160
2821 GCGCTACTAGGCTTGGCAGGTGGAATGGCACTTATCATATGACCACTCACCCCTATG 2880
3161 GCGCTACTAGGCTTGGCAGGTGGAATGGCACTTATCATATGACCACTCACCCCTATG 3220
2881 TCGGATTTGGGCTCTAGTGGCTGCGGAGACTGCGGAGTGGCGCTTGAAGCTTATCTTCC 2940
3221 TCGGATTTGGGCTCTAGTGGCTGCGGAGACTGCGGAGTGGCGCTTGAAGCTTATCTTCC 3280
2941 AGTCCGATGAGAAAGATCATGTCTGAGGAGCGGAGACAGCTGCTTGTGGGAGCATTT 3000
3281 AGTCCGATGAGAAAGATCATGTCTGAGGAGCGGAGACAGCTGCTTGTGGGAGCATTT 3340
3001 TTACAGGACCTTCCGCTGTCGCGGACATTTGGCGGAGGATCTCTTGTGGCCACGCTAT 3060
3341 TTACAGGACCTTCCGCTGTCGCGGACATTTGGCGGAGGATCTCTTGTGGCCACGCTAT 3400
3061 GGCCTATACCTCAAGGGGTGAGTCTTCTGCGCCCATCATCTGTTACCGCCAGAGACA 3120
3401 GGCCTATACCTCAAGGGGTGAGTCTTCTGCGCCCATCATCTGTTACCGCCAGAGACA 3460
3121 CGTGGCTTTTGGGACCATATGATGTGAGCATGACGGGGCGGACAAAGACAGAGCT 3180
3461 CGTGGCTTTTGGGACCATATGATGTGAGCATGACGGGGCGGACAAAGACAGAGCT 3520

QY 3181 GGGGAAATTCAGGCTCTGTCCACAGTCACTAGTCTTCTCGGAACATCCATCTCGGG 3240
DB 3521 GGGGAGATTCAGGCTCTGTCCACAGTCACTAGTCTTCTCGGAACATCCATCTCGGG 3580
QY 3241 GTTTTGTGAGCTGTCTACATGAGCTGGCAACAAGCTGTGGCCGGCTACGGGGTCCG 3300
DB 3581 GTCTTATGAGCTGTCTACATGAGCTGGCAACAAGCTGTGGCCGGCTACGGGGTCCG 3540
QY 3301 GTCAAGAGATGTACTCCAGTCTGAGGGGACTTATGAGGTGGCCAGCCCGCTGG 3360
DB 3641 GTCAAGAGATGTACTCCAGTCTGAGGGGACTTATGAGGTGGCCAGCCCGCTGG 3700
QY 3361 ACTAATCTTTGAGAGCCGTGACGTGTGAGAGCGGTGACCTGTACCTGTACAGCGGAC 3420
DB 3701 ACCAAATCTTTGAGAGCCGTGACGTGTGAGAGCGGTGACCTGTACCTGTACAGCGGAC 3760
QY 3421 GCTGATGATCCCGGCTCGAAGAGCGGGGACAAAGGGGAGCGCTACCTCCCGAGA 3480
DB 3761 GCTGATGATCCCGGCTCGAAGAGCGGGGACAAAGGGGAGCGCTACCTCCCGAGA 3820
QY 3481 CCTCTTTCACCTTGAAGGGGTCTCAGAGAGCCGGTGTATGCCCCAGGGGACAGCT 3540
DB 3821 CCTCTTTCACCTTGAAGGGGTCTCAGAGAGCCGGTGTATGCCCCAGGGGACAGCT 3880
QY 3541 GTGAGAGTCTTTCGGGAGAGCTGTGTGCTCTTCGGGGCGTGGCTAAATGTCATAGATTTGATC 3600
DB 3881 GTGAGAGTCTTTCGGGAGAGCTGTGTGCTCTTCGGGGCGTGGCTAAATGTCATAGATTTGATC 3940
QY 3601 CCCGTTAGACACTTCGACATCGTGAACCGCGTCCCGCCACTTATGATGACACAGACCA 3660
DB 3941 CCCGTTAGACACTTCGACATCGTGAACCGCGTCCCGCCACTTATGATGACACAGACCA 4000
QY 3661 CCTGCTGTGCCCCAGACCTATCAGAGTGGGACTTGTATGATGCCCCAGCTGGAGTGAAG 3720
DB 4001 CTTGCTGTGCCCCAGACCTATCAGAGTGGGACTTGTATGATGCCCCAGCTGGAGTGAAG 4060
QY 3721 AGACCAAAAGTCTGTGCGATATGCTGTCAAGGGGTATAAAGTCTAAGTCTTAATCCC 3780
DB 4061 AGACCAAAAGTCTGTGCGATATGCTGTCAAGGGGTATAAAGTCTAAGTCTTAATCCC 4120
QY 3781 TCAATGCTGTGCAACCTGGGGTTTGGGGCTGACTTGTCTAAGGACATGATCAATCAATCCC 3840
DB 4121 TCGGTGTGCTGCAACCTGGGGTTTGGGGCTGACTTGTCTAAGGACATGATCAATCAATCCC 4180
QY 3841 AACATTAGAGCTGAGAGTCAAGACTGTGACGACCGGGGCGCCATCAGTACTCCACATAT 3900
DB 4181 AACATTAGAGCTGAGAGTCAAGACTGTGACGACCGGGGCGCCATCAGTACTCCACATAT 4240
QY 3901 GGGCAATTCCTGCGGATGGGGGCTGTGCGGGCGGCGCTTACGACATCATATGTGAT 3960
DB 4241 GGGCAATTCCTGCGGATGGGGGCTGTGCGGGCGGCGCTTACGACATCATATGTGAT 4300
QY 3961 GAAAGGCAATGCGGAGCTCTACACCACTCTTGGACATCGGAAGAAGTCTTGAACAACA 4020
DB 4301 GAAAGGCAATGCGGAGCTCTACACCACTCTTGGACATCGGAAGAAGTCTTGAACAACA 4360
QY 4021 GAGACAGCTGGGGTCAAGACTGATGTGCTGACTACAGTACGCCCCCTGGGTCACTGACA 4080
DB 4361 GAGACAGCTGGGGTCAAGACTGATGTGCTGACTACAGTACGCCCCCTGGGTCACTGACA 4420
QY 4081 ACCCCCCCAACCATAGAGAGGTGCGCTTGGGAGAGAGGCGAGATCCCTTTCTAT 4140
DB 4421 ACCCCCCCAACCATAGAGAGGTGCGCTTGGGAGAGAGGCGAGATCCCTTTCTAT 4480
QY 4141 GGGAGGGGAGATTCCTCTGTCTTATACATCAAGGAGAGAAACATCTGATCTTCTGCAATTA 4200
DB 4481 GGGAGGGGAGATTCCTCTGTCTTATACATCAAGGAGAGAAACATCTGATCTTCTGCAATTA 4540
QY 4201 AAGAAAAAGTGTGAGAGCTGCGGGCGGCTTTCGGGGTATGGCTTGAATCAGTGGCA 4260
DB 4541 AAGAAAAAGTGTGAGAGCTGCGGGCGGCTTTCGGGGTATGGCTTGAATCAGTGGCA 4600

QY 4261 TACTACAGAGGTTGAGAGCTCTCCGTAATACCACTCAGGAGAGAGTGTGCTCGCC 4320
DB 4601 TACTACAGAGGTTGAGAGCTCTCCGTAATACCACTCAGGAGAGAGTGTGCTCGCC 4660
QY 4321 ACCGAGCCCTCATGACAGGGATATCTGGGGACTTGTGATCCGTATGACGTGACAGTGA 4380
DB 4661 ACCGAGCCCTCATGACAGGGATATCTGGGGACTTGTGATCCGTATGACGTGACAGTGA 4720
QY 4381 GGGTCACTCAAGTTGTAGACTTCACTTGTGAGAGCCCACTTACATCAACATCAACAGATT 4440
DB 4721 GGGTCACTCAAGTTGTAGACTTCACTTGTGAGAGCCCACTTACATCAACATCAACAGATT 4780
QY 4441 GTCCCTCAAGAGAGCTGTCTCAGGTAGCCAGCCCGGGGTTCGACAGGGATAGGGGAAACTG 4500
DB 4781 GTCCCTCAAGAGAGCTGTCTCAGGTAGCCAGCCCGGGGTTCGACAGGGGAAACTG 4840
QY 4501 GGCATTATAGGATATGTTTCACTGCTGAGCGAGCTCAGGAATGTTTGAAGTGAAGTG 4560
DB 4841 GGCATTATAGGATATGTTTCACTGCTGAGCGAGCTCAGGAATGTTTGAAGTGAAGTG 4900
QY 4561 CTCTGAGTGTCTACGACGAGGGGCGGATGATGAGCTGACACATCGAGACACACC 4620
DB 4901 CTCTGAGTGTCTACGAGGAGGCGGATGATGAGCTGACACATCGAGACACACC 4960
QY 4621 GTCAAGCTCAGAGCGTATTTCAACACGCCCCGTTTGTCTGTGTCCCAACATCTTGA 4680
DB 4961 GTCAAGCTCAGAGCGTATTTCAACACCTGGTTTGTCTGTGTCCCAACATCTTGA 5020
QY 4681 TTTTGGGAGGCAAGTTTTCACCGGCTCAGACACATGATGAGCCACTTCTTCCCAACA 4740
DB 5021 TTTTGGGAGGCAAGTTTTCACCGGCTCAGACACATGATGAGCCACTTCTTCCCAACA 5080
QY 4741 AAGCAATCGGGGAAAAATTTGCACTTAAAGCTTACCAAGCTTACAGTGTGCTGATG 4800
DB 5081 AAGCAATCGGGGAAAAATTTGCACTTAAAGCTTACCAAGCTTACAGTGTGCTGATG 5140
QY 4801 GCCAAAGCCCCCCCCGTCGTGGGACGTCAATGTGGAAGTGTGACTGACTCAAGCCC 4860
DB 5141 GCCAAAGCCCCCCCCGTCGTGGGACGTCAATGTGGAAGTGTGACTGACTCAAGCCC 5200
QY 4861 ACACGTGTGGGCCCCCACCTCTCTGTACCGGTGGGCTGTGTTCAACAGAGTCAAC 4920
DB 5201 ACACGTGTGGGCCCCCACCTCTCTGTACCGGTGGGCTGTGTTCAACAGAGTCAAC 5260
QY 4921 CTCACACATCCCGTGAAGAAATACATGCGCACCTGATGCAAGCCGACTTGAAGTCAATG 4980
DB 5261 CTCACACATCCCGTGAAGAAATACATGCGCACCTGATGCAAGCCGACTTGAAGTCAATG 5320
QY 4981 ACCAGACATGGGTCTTGGCAGGGGAGTCTTGGCGGCGGCTGCGCGCTATGCTCGGG 5040
DB 5321 ACCAGACATGGGTCTTGGCAGGGGAGTCTTGGCGGCGGCTGCGCGCTATGCTCGGG 5380
QY 5041 ACCGGGTGTGTTGATCATCGGCGCTTGACATTAACAGAGGACCGTGTGGCGG 5100
DB 5381 ACCGGGTGTGTTGATCATCGGCGCTTGACATTAACAGAGGACCGTGTGGCGG 5440
QY 5101 GACAAAGGTCTCTTATGAGAGCTTGTGATGAGTGAAGAAATGTGCTCTAGGGCGGCT 5160
DB 5441 GACAAAGGTCTCTTATGAGAGCTTGTGATGAGTGAAGAAATGTGCTCTAGAGCGGCT 5500
QY 5161 CTGATTGAAGAGGGGACAGGAGATGCGGAGATGCTGAAGTCCAAATCCAAAGCTTATG 5220
DB 5501 CTGATTGAAGAGGGGACAGGAGATGCGGAGATGCTGAAGTCCAAATCCAAAGCTTATG 5560
QY 5221 CAGCAAGCTTCCAAACAGGCTCAAGCATACCAACCACTGTGACAGGCTTCAATGGCCAG 5280
DB 5561 CAGCAAGCTTCCAAACAGGCTCAAGCATACCAACCACTGTGACAGGCTTCAATGGCCAG 5620
QY 5281 GTAGAACAAATTTCTGGGCAAAACATATGTGAACTTATTAAGCGGATCCAAATCCTGCA 5340
DB 5621 GTAGAACAAATTTCTGGGCAAAACATATGTGAACTTATTAAGCGGATCCAAATCCTGCA 5680
QY 5341 GGACTATCAACCTGCGAGGAAACCTGCAAGTATGCTTCCATGATGCGTTCAAGTCCGCC 5400

Db 5681 GGAATCAACACCTGCGAGGAAACCTGCTGAGCTTCCATGATGCAATTCAGTCCGCC 5740
Qy 5401 CTGACAGTCCGCTGTCAACAGACCACTATCTTCTCAACATTTTGGGGGGCTGAGCTA 5460
Db 5741 CTGACAGTCCGCTGTCAACAGACCACTATCTTCTCAACATTTTGGGGGGCTGAGCTA 5800
Qy 5461 GCAATCCAAATTTGACACACCGCGGGGGCCACTGGCTTCGTTGCAATGTCGCTTAATGGA 5520
Db 5801 GCAATCCAAATTTGCGCTCCCGCGGGGGCTACCGGCTTCGTCGCAATGTCGCTGGTGG 5860
Qy 5521 GCTGCGGTAGGAGATATAGGCTTATAGTAAAGTGTAGTGAATCTGCGAGGATATAGT 5580
Db 5861 GCTGCGGTAGGAGATATAGGCTTATAGTAAAGTGTAGTGAATCTGCGAGGATATAGT 5920
Qy 5581 GCGGGCAATTTGCGGGGGCTCTGTCGCAATTCAGATCATGTCGCGAGAAACCTTCATG 5640
Db 5921 GCGGGCAATTTGCGGGGGCTCTGTCGCAATTCAGATCATGTCGCGAGAAACCTTCATG 5980
Qy 5641 GAGGATGTCGTCATCTTGCTGCTGAGAAATCTGTCCTCGGGTGCCTTGATAGTGGAGTC 5700
Db 5981 GAGGATGTCGTCATCTTGCTGCTGAGAAATCTGTCCTCGGGTGCCTTGATAGTGGAGTC 6040
Qy 5701 ATCTGCGGGGCAATCTGCGCCGACAGTGGGACCGGGGGAAGGCGCTGTCATATGATG 5760
Db 6041 ATCTGCGGGGCAATCTGCGCCGACAGTGGGACCGGGGGAAGGCGCTGTCATATGATG 6100
Qy 5761 AATAGACTCATTTGCTTGTCTCCAGAGAAATCATGTCGCCCCCACCCTACATGATGAG 5820
Db 6101 AATAGACTCATTTGCTTGTCTCCAGAGAAATCATGTCGCCCCCACCCTACATGATGAG 6160
Qy 5821 GAGTCCGATGCTGTGCGAGCGTGTGACCCACTATCTTGAGCTCCCTTACCATTAACAGCTG 5880
Db 6161 GAGTCCGATGCTGTGCGAGCGTGTGACCCACTATCTTGAGCTCCCTTACCATTAACAGCTG 6220
Qy 5881 CTCGAGAGACTCCACACTGGAATTAAGAGACTGCCCCATCCCATGCAATGAGCTGAG 5940
Db 6221 CTCGAGAGACTCCACACTGGAATTAAGAGACTGCCCCATCCCATGCAATGAGCTGAG 6280
Qy 5941 CTCGCGATGTGTGGAGCTGGGTTTGCACCATCTCTAAACAGCTTTAAATAATTTGGCTAGC 6000
Db 6281 CTCGCGATGTGTGGAGCTGGGTTTGCACCATCTCTAAACAGCTTTAAATAATTTGGCTAGC 6340
Qy 6001 TCCAAATTAATTTCCAAAGATGCCCCGCTTCCTTGTCTCCCTCAAAAAGGGGTAAAG 6060
Db 6341 TCCAAATTAATTTCCAAAGATGCCCCGCTTCCTTGTCTCCCTCAAAAAGGGGTAAAG 6400
Qy 6061 GCGGTGTGGGCGGACCTGAGCATATGACCAACAGTGTCTTTGCGGCGCAATATCTCT 6120
Db 6401 GCGGTGTGGGCGGACCTGAGCATATGACCAACAGTGTCTTTGCGGCGCAATATCTCT 6460
Qy 6121 GGCATGTGCGCTTGGGCTCATAGAGATACGGGGGCTTAAGACTGTCATGAATATCTGG 6180
Db 6461 GGCATGTGCGCTTGGGCTCATAGAGATTAAGGGGGCCAAACCTGCAATGAATATCTGG 6520
Qy 6181 CAGGGGACTTTTCTTCAATTTGTTAAACGGAAGGCGAGTGCCTGCGGAAACCGGACCA 6240
Db 6521 CAGGGGACTTTTCTTCAATTTGTTAAACGGAAGGCGAGTGCCTGCGGAAACCGGACCA 6580
Qy 6241 AACTTTAAGTGTGCGCATCTGGAAGGCTGAGGCTCAAGTACGCGAGAGTGAACGACAC 6300
Db 6581 AACTTTAAGTGTGCGCATCTGGAAGGCTGAGGCTCAAGTACGCGAGAGTGAACGACAC 6640
Qy 6301 GGGTCATACCACTATACAGAGACTCAACCATGATTAAGTAAAGTCCCTGCGCACTA 6360
Db 6641 GGGTCATACCACTATACAGAGACTTAACCATGATTAAGTAAAGTCCCTGCGCACTA 6700
Qy 6361 CCGCTCCGAGTCTTTCTGCGGTGGAAGGATGAGATCCATAGTTTGGCCGACCA 6420
Db 6701 CCGCTCCGAGTCTTTCTGCGGTGGAAGGATGAGATCCATAGTTTGGCCGACCA 6760
Qy 6421 CCGAAGCGTTTCTGCGGATGAGGCTGCTTCTGCGGTGGGCTTAAATTCATTTGTGCTC 6480

Db 6761 CCGAAGCGTTTCTGCGGATGAGGCTGCTGCTGCGGTGGGCTTAATTCATTTGTGCTC 6820
Qy 6481 GGGTCCGAGCTTCTTGGGACCTTGAAACCGACAGAGTATTTGATGTCCATGCTTACA 6540
Db 6821 GGGTCTCAGCTCTCTTGGCGATCTCTGAACCTGACACAGACGTATTTAGTCCATGCTTACA 6880
Qy 6541 GATCATCTCATATACAGGCGGAGACTGACGCGGGGTTTACGCGGGGGTCAACCCCA 6600
Db 6881 GATCATCTCATATACAGGCGGAGACTGACGCGGGGTTTACGCGGGGGTCAACCCCG 6940
Qy 6601 TCCGAGCAAGCTCTTCCGAGGACCACTATCCGACCATCGCTGCGAGCACTGACCC 6660
Db 6941 TCCGAGCAAGCTCTTCCGAGGACCACTATCCGACCATCGCTGCGAGCACTGACCC 7000
Qy 6661 ACCGACGGCAAGCTTATGATGTGAGACATGCTGATGCTTAACTGTTCATGCGGGGAGAT 6720
Db 7001 ACCGACGGCAAGCTTATGATGTGAGACATGCTGATGCTTAACTGTTCATGCGGGGAGAT 7060
Qy 6721 GTGACTCGATGATGAGTCTGGGTCGAAAGTGGTGGTCTGAGCTCTCGAACCCAAATGCTC 6780
Db 7061 GTGACTCGATGATGAGTCTGGGTCGAAAGTGGTGGTCTGAGCTCTCGAACCCAAATGCTC 7120
Qy 6781 GAAGAAAGAGCGACCTTGAAGCTTCGATACCATCAGAAATACATGCTCCCAAGAGAG 6840
Db 7121 GAAGAAAGAGCGACCTTGAAGCTTCGATACCATCAGAAATACATGCTCCCAAGAGAG 7180
Qy 6841 TTCCCAACAGCTTTACCGGCTTGGGCAACGCGCTGATTAACACCAACCGCTTTGTGGAATCG 6900
Db 7181 TTCCCAACAGCTTTACCGGCTTGGGCAACGCGCTGATTAACACCAACCGCTTTGTGGAATCG 7240
Qy 6901 TGGAAAGGCGAGATTAACCAACGCGGACATGTTGCGGGGCTGTGCTCTCCCTCCCTAGG 6960
Db 7241 TGGAAAGGCGAGATTAACCAACGCGGACATGTTGCGGGGCTGTGCTCTCCCTCCCTAGG 7300
Qy 6961 AAAACCCGACGCGCTCCCAAGAGAGCGCGGACAGTGGGCTTGAAGTGAAGCTCCATA 7020
Db 7301 AAAACCCGACGCGCTCCCAAGAGAGCGCGGACAGTGGGCTTGAAGTGAAGCTCCATA 7360
Qy 7021 GAGATGCTCTTCAACAGCTGCGCATTAATGCTTTTGGCCAGGCCCCCCCAAGCGGAGAT 7080
Db 7361 GAGATGCTCTTCAACAGCTGCGCATTAATGCTTTTGGCCAGGCCCCCCCAAGCGGAGAT 7420
Qy 7081 TCAAGGCTTTTCAACGAGGAGGCGGCGCTGCGGATTCGCGGAGTGAAGCTCTGATGAG 7140
Db 7421 TCAAGGCTTTTCAACGAGGAGGCGGCGCTGCGGATTCGCGGAGTGAAGCTCTGATGAG 7480
Qy 7141 TTGGCCCTTTGAGAGACAGGTTCCATCTCTTCCATGCCCCCTCGAGGGGAGACTTGA 7200
Db 7481 TTGGCCCTTTGAGAGACAGGTTCCATCTCTTCCATGCCCCCTCGAGGGGAGACTTGA 7540
Qy 7201 GATTCAGACCTTGAAGCTTGAAGCTGAGTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 7260
Db 7541 GATTCAGACCTTGAAGCTTGAAGCTGAGTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 7600
Qy 7261 GCTCCGCGGCTCGGAGCTCGGGGCTCTGCTTCACTTGTGCTCGAGGAGAGAGCATCGTCTGTG 7320
Db 7601 ACCCGGCGCTCAGGCTCGGGGCTTGGTCTTACTTGTGCTCGAGGAGAGAGCATCGTCTGTG 7660
Qy 7321 TGTGCTCTCATATGCTCTTGAAGCGGGCTTAAATACCTCTTGTAGTCCGGAAGAG 7380
Db 7661 TGTGCTCTCATATGCTCTTGAAGCGGGCTTAAATACCTCTTGTAGTCCGGAAGAG 7720
Qy 7381 GAGAAGTTACCAATTAACCTTGAAGCAATCCCTGTGAGATATCAACCAAGGTATAC 7440
Db 7721 GAAAAGTTGCAATTTGAGGCTTGAAGCAATCCCTGTGAGATATCAACCAAGGTATAC 7780
Qy 7441 TGTACCAACCAAGAGCGGCTCACTAAGGGCTTAAAGGTAACTTTTGAATAGATGCA 7500
Db 7781 TGTACCAACCAAGAGCGGCTCACTAAGGGCTTAAAGGTAACTTTTGAATAGATGCA 7840
Qy 7501 GTGCTGAGCTCTTACATGAGCTGATCTTAAAGAGCAATTAAGCTAGCGGCTCCAGAGTC 7560
Db 7841 GCGCTGAGCGCTCATTAATGAGCTGATCTTGAAGAGCAATTAAGCTAGCGGCTCCAGAGTC 7900

TELEX: MWI 64470
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 9511 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: single
: TOPOLOGY: linear
US-07-925-695-7

Query Match 62.1%; Score 5653.8; DB 1; Length 9511;
Best Local Similarity 75.5%; Pred. No. 0;
Matches 6868; Conservative 104; Mismatches 2127; Indels 0; Gaps 0;

```
Qy 1 ATGAGCAAAATCTTAACCTCAAGAAAAACAAAAGAAACACCAACCGTCGCCCAAA 60
Db 342 ATAGAGCAAAATCTTAACCTCAAGAAAAACAAAAGAAACACCAACCGTCGCCCAAG 401
Qy 61 GAGCTTAAGTTTCCGGGGGCGGCGGAGATCGTTGGCGGAGTAACTTGTGGCCGCGAGG 120
Db 402 GAGCTTAAGTTTCCGGGGGCGGCGGAGATCGTTGGCGGAGTAACTTGTGGCCGCGAGG 461
Qy 121 GGGCCCAAGTTGGGTGGTGGCGGCAAGAGAACTTCGAGCGGTCCCAAGCCAGTGA 180
Db 462 GGGCCCAAGTTGGGTGGTGGCGGCAAGAGAACTTCGAGCGATCCAGCCGCGTGA 521
Qy 181 AGGCGGCAAGCCCAATCCCTTAAGATCGGCGCTCCACTGGCAAAATCTGGGGAAAAACAGGA 240
Db 522 CGAGCGCAAGCCCAATCCCTTAAGATCGGCGCTCCACTGGCAAAATCTGGGGAAAAACAGGA 581
Qy 241 TACCCCTGGCCCTTAAACGGGAATGAGGAACTCGGCTGGCGAGAGATGGCTCTGTCCCC 300
Db 582 TATCTTGGCCCTTAAACGGGAATGAGGAACTCGGCTGGCGAGAGATGGCTCTGTCCCC 641
Qy 301 CGAGTTCCCGTCCCTCTTGGGGGCCCCAATGACCCCGGCAATAGTGGCGCAAGTGGGT 360
Db 642 CGGGGCTCTGTCTTCTTACCTTGGGGGCCCCAAGCCCGGCAATAGTGGCGCAATAGTGGGT 701
Qy 361 AAGGTCAATGATACCTTAACGATGCGCTTTGCGGAACTCAATGGGGTCAATCCCTGTGCTG 420
Db 702 AAGGTCAATGATACCTTAACGATGCGCTTTGCGGAACTCAATGGGGTCAATCCCTGTGCTG 761
Qy 421 GGGGCCCCGCTCGGCGGCGTCCAGAGCTCTCGCGCATGGCGTGAAGTCTTGAAGAC 480
Db 762 GGGGCCCCGCTCGGCGGCGTCCAGAGCTCTCGCGCATGGCGTGAAGTCTTGAAGAC 821
Qy 481 GGGGTAAATTTTGGCAAGGAACTTACCGGTTGCTCTTTTCTATCTTGTGCTGGC 540
Db 822 GGGGTAAATTTTGGCAAGGAACTTACCGGTTGCTCTTTTCTATCTTGTGCTGGC 881
Qy 541 CTGCTGCTGATCAACCAACCGGCTCGCTGCGGAGTGAAGTGAAGATCAGTACCGGC 600
Db 882 CTGCTGCTGATCAACCAACCGGCTCGCTGCGGAGTGAAGTGAAGATCAGTACCGGC 941
Qy 601 TACATGATGATCAAGACTGCAACCAATGAGAGATTAATCTGGAGCTCCAGGCTGTGTC 660
Db 942 TACATGATGATCAAGACTGCAACCAATGAGAGATTAATCTGGAGCTCCAGGCTGTGTC 1001
Qy 661 CTTCAGTCCCCGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 720
Db 1002 CTTCAGTCCCCGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1061
Qy 721 CCGGCTCAACGAGATGCGGCTGCGAGCGGCGCGGCGGCGGCTCAACGAGGCTTGGCGAGC 780
Db 1062 CAAGTAAACCCCAACGAGATGCGGCTGCGAGCGGCGGCGGCGGCTCAACGAGGCTTGGCGAGC 1121
Qy 781 CACATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Db 1122 CACATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1181
Qy 841 TGGGTTGGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 900
Db 1182 TGGGTTGGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1241
```

```
Qy 901 GTCCAGAGTCAATGCTCCATCTACCTGATACATCACTGAGACACCGCATGGATGG 960
Db 1242 ACCCAAGAGTGCACACTTTCATCTTACCAAGTCACTACCGGCATTCGATGGATGG 1301
Qy 961 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Db 1302 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1361
Qy 1021 CCGGAGTCAATTAATGATCACTTAAGGAGGCTCATTTGGGGCGGTCATGCTTGGCC 1080
Db 1362 CCGGAGTCAATTAATGATCACTTAAGGAGGCTCATTTGGGGCGGTCATGCTTGGCC 1421
Qy 1081 TACTTCTTAATGAGAGGAGCGTGGGCGAAAGTGTGTCATCTTCTGTGGCGCGGG 1140
Db 1422 TACTTCTTAATGAGAGGAGCGTGGGCGAAAGTGTGTCATCTTCTGTGGCGCGGG 1481
Qy 1141 GTGAGCGGCGCACCCATCTGTTGGGGGTTCTGCGCGGAGACACCGGCGCGCTCAC 1200
Db 1482 GTGAGCGGCGCACCCATCTGTTGGGGGTTCTGCGCGGAGACACCGGCGCGCTCAC 1541
Qy 1201 AGCTTAATTAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
Db 1542 AGCTTAATTAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1601
Qy 1261 CACATCAACCGCACCGGCTGAATGATGATGATGATGATGATGATGATGATGATGATG 1320
Db 1602 CACATCAACCGCACCGGCTGAATGATGATGATGATGATGATGATGATGATGATGATG 1661
Qy 1321 CTGTTCTACACCGACAGCTTCAATCTGTCAGATGTCGGAACGATGTCGCTGCGC 1380
Db 1662 CTGTTCTACACCGACAGCTTCAATCTGTCAGATGTCGGAACGATGTCGCTGCGC 1721
Qy 1381 AGTATGAGGAGGCTTCCGGGAGGAGTGGGGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
Db 1722 AGTATGAGGAGGCTTCCGGGAGGAGTGGGGGAGGAGGAGGAGGAGGAGGAGGAGG 1781
Qy 1441 CCAGAGGATTAATGAGAGGCTTATGCTGAGCACTACCAACCAAGGAGTGGCGTGTCC 1500
Db 1782 GATGAGGATTAATGAGAGGCTTATGCTGAGCACTACCAACCAAGGAGTGGCGTGTCC 1841
Qy 1501 GCGAGAGCTGTGTGTGGGCGGAGTGTATCTGTTACCCCGGAGGCTGAGTGGGAGC 1560
Db 1842 GCGAGAGCTGTGTGTGGGCGGAGTGTATCTGTTACCCCGGAGGCTGAGTGGGAGC 1901
Qy 1561 ACCGACAGGCTTGGAGGCGGCTCACTTACAGTGGGGGAGGAGGAGGAGGAGGAGG 1620
Db 1902 ACCGACAGGCTTGGAGGCGGCTCACTTACAGTGGGGGAGGAGGAGGAGGAGGAGG 1961
Qy 1621 TTGAACAGCACTGACCAACCGGCTGGGGTCAATGCTGCGGCTGACGATGAGTCTTCT 1680
Db 1962 TTGAACAGCACTGACCAACCGGCTGGGGTCAATGCTGCGGCTGACGATGAGTCTTCT 2021
Qy 1681 GGGTACACCAAGACTTGGCGGCGGAGCACCGCTGCGGATCTAGAGCTTCAACCGCAGC 1740
Db 2022 GGGTACACCAAGACTTGGCGGCGGAGCACCGCTGCGGATCTAGAGCTTCAACCGCAGC 2081
Qy 1741 ACGGACCTGTTGGGCGGCGGAGCTGTTTGAAGACATCTGATACCACTTCAACCA 1800
Db 2082 ACGGACCTGTTGGGCGGCGGAGCTGTTTGAAGACATCTGATACCACTTCAACCA 2141
Qy 1801 TGGGCTTGGGCGGCTTGGCTGACAGCAAGTGGCTGATGATGATGATGATGATGATGATG 1860
Db 2142 TGGGCTTGGGCGGCTTGGCTGACAGCAAGTGGCTGATGATGATGATGATGATGATGATG 2201
Qy 1861 CATTAACCTTGCACAGTAACTAATACATCTTCAAAATTAAGATGATGATGATGATG 1920
Db 2202 CATTAACCTTGCACAGTAACTAATACATCTTCAAAATTAAGATGATGATGATGATGATG 2261
Qy 1921 GAGCAGAGGCTCAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1980
Db 2262 GAGCAGAGGCTCAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2321
Qy 1981 AGAGACAGAGTCACTGCTCTTGTGTAATCCACACGAGATGGGCAATTTTAACT 2040
```

Db 2222 AGGATATGGGGYCGAGCAGAGTCCACTGCTGATTTCCACTAGTGAGGGGGTGTCCCA 2281
Oy 2441 TGCTCTACTCGGACCTGCCCGCTTGACATGAGTCTTCCCACTCCCAACCAAAATATC 2100
Db 2382 TGCTCTCTCTGACCTTACAGCAGCTATCCATGGCTTATGACCTCCACCAAAATATC 2441
Oy 2101 GTGACGTACAAATTCATGATAGGCTTATCACTGCTCCCTCAAAAATATCATGTCGAGT 2160
Db 2442 GTGACGTGACAGTACCTTATACGGAATTTCTCCGGCTGACCAAAATATCATGTCGAGT 2501
Oy 2161 GAGTGGGTATATCTTATCTTCTGCTCTTACGGAACGCGAGGTTTCCGCTGTATG 2220
Db 2502 GAGTGGGTATATCTTCTTCTTCTTGTGTGACAGCCAGGRTCTGTGATGCTTTTG 2561
Oy 2221 ATGCTCATCTTGTGGGCGAGGCGCAAGCAGCAGTACAGAAAGTGTGATCTTGACGCT 2280
Db 2562 ATGCTCAACATATCTGGGCGCAAGCGGAGCGGCTTGAGAAAGTCTATCATCTTGACCTC 2621
Oy 2281 GCGAGCGAGCTAGCTGCAATGAGCTTCTCATATTTTGTATCTTTTGTGTGCTGCTTG 2340
Db 2622 GCTAGTGTGCTAGGCAATGATCCGCTGTGTGTTTTCATCTTTTACAGCGGCTGG 2681
Oy 2341 TACATCAAGGCTGGGTATGTCCTCTTACGTAATCTTATCTTCACTGGCTGTGCTTT 2400
Db 2682 TACTTAAAGGCGAGGAGTGTCCCGTGGCCAGTACTGTBCTCGGCTGTGTGCTTTC 2741
Oy 2401 AGCTTACGCTAGCTAGCATTGCCCCAAGAGCTTATGCTTATGACGATCTGTGATGCG 2460
Db 2742 CTCTCTCTAGTCTGCTGCTTTTACACAGCAGGCTTATGCTTGAGCGCTGTAAAGGG 2801
Oy 2461 CAGATAGAGCGGCTGCTGCTGTATGATCACTCTTTTACTCTTACCCCCGGGTATAG 2520
Db 2802 GAATCGGGGCTGGCCATATATGATATATATCAATCTTTTACTTTTACCCGAGTACAG 2861
Oy 2521 ACCCTTCTCAGCCGCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
Db 2862 ATCTCTCAGAGCCGTTTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2921
Oy 2581 GTCCAGAGTGGGCGACCACTATGACAGGTGCGCGGTGCGGTGATGATCATATGAGGCC 2640
Db 2922 ATTACGACATGGGTTCCCCCTGAGGTCCGAGGGGGGCGGTGACGGGATATCTGGGTG 2981
Oy 2641 GTCCGCAATTTCTACCCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2700
Db 2982 GCTGATTTCTAACCCAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3041
Oy 2701 GGGCTGTGTATCTTCTTAAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2760
Db 3042 GGGCTGTGTATCTTCTTAAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3101
Oy 2761 GCTCTACTGAGGATGTGACCATGACAGGATCTCGGCGGGGGAGGATGTCAGATG 2820
Db 3102 GCTTTGCTACAGATGTGTATCCCTGTGTGAAAACCTCGGCGGGGCTAGGTATCATCAAGT 3161
Oy 2821 GCGCTACTAGCCCTTGTGAGGTGTGACTGCACTTATGATGACCACTTCACTCATG 2880
Db 3162 CTGTTTATACCATATGAGGATGTGACCGGCACTTATCATCTAGACCACTCTCCCTTTA 3221
Oy 2881 TCGATTTGGGCTGTAGTGTGCTGTGGGACTCGGCGGTGCGCTGTAGGCTTATCATCTTC 2940
Db 3222 TCAACTTGGGCGGCCAGGGTGTTRCGGGACCTGGGCAATCGCCGTGTAGACCTGTGTGTTC 3281
Oy 2941 AGTCGATGAGAAAGAAATGCTATGTCTGGGGAGCGGAGACAGCTGTGTGGGAACTT 3000
Db 3282 AGCCCAATGAGAAAGAAATGCTATGTGTGGGGGCTGTGAGACAGTGTGTGAGACATC 3341
Oy 3001 TTACACGACTTCCCGTGTCCGCGCCGACTTGTCTGGAGGTCTTCTTGTGGCCAGCTAT 3060
Db 3342 CTGATGTGCTTCCCGGTCTCCGCGAGGCTAGGTAGGATGTGTGTGTGTGTGTGTGTGTGTGT 3401
Oy 3061 GGCTTACTTCCAGGGGTGAGTCTTCTCGCCCCCATCACTGCTTACGCCAGCAGACA 3120

Db 3402 GGCTACACTCCAGGGGTGAAGACTCTAGTCTCCCATTTACTGCTTACACTGACAAACT 3461
Oy 3121 CGTGGCTTTTGGGACCATATGATGTGACATGACGGGGCGGAGACAGAAAGGCT 3180
Db 3462 CGTGTCTCTGTGGTGTCTATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3521
Oy 3181 GGGGAAATTCAGGTCTGTGTCAAGTCACTCACTGCTTCTCTGGAACATCTTCGGGG 3240
Db 3522 GGGCAGGTCCAGATCTGTCTTCCGTCACAAACTTCTTGGGGAATCTCAATTCGAGC 3581
Oy 3241 GTTTTGTGACGTCTACCATGAGGTGCGCAAAAGACTCTGGCGGGCTCACGGGGTCCG 3300
Db 3582 GTCTCTGACAGTATATACCGGGCTGTGTATTAAGACTTGGCGGCCCAAGGAGCA 3641
Oy 3301 GTACACAGATGTACTCGATGCTGAGGGGAGCTTATAGGGTGTGCCAGCCCTGTGG 3360
Db 3642 GTTACTCAGATGTATACCGAGCGGAGAGGGGAGCTGTGTGGATGTGTGTGTGTGTGTGTGT 3701
Oy 3361 ACTAAATCTTTGAGGCGGTGACAGTGTGAGCGGTGTGACCTGTACTGTGTCAAGCGA 3420
Db 3702 ACTAAGTCAATTGAGCCCTGTACTGTGCGGGCGGTAGACTCTTACTGTGTCAAGCGA 3761
Oy 3421 GCTGATGTATCCCGCTGGAAGCGGGGCAAAAGGGGAGCGCTACTCTCCCGAGA 3480
Db 3762 GTGATGTATTCGGTCCGAGGAAAGATGACCGAGGGGTGTGATTAATCTTGTGCGAAG 3821
Oy 3481 CCTCTTCAACTTGAAGGGGTCTTCAAGAGGCGCGGTGTATGCCCCAGGGCCAGCT 3540
Db 3822 CCGCTCTCAACCTCAAGAGATATTCGGAGGGGCCGTGTCTGTCTGTGTGTGTGTGTGTGTGT 3881
Oy 3541 GTGAGTCTTCCGGGAGCTGTGTGTCTTCCGGCGGTGTGTGTGTGTGTGTGTGTGTGTGT 3600
Db 3882 GTGGGCTGTGTGAGAGCGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3941
Oy 3601 CCGGTGAGACACTGTGACATGTGTGACGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3660
Db 3942 CCGGTGATCATCTGTATGT 4001
Oy 3661 CCGT 3720
Db 4002 CGAGCTGT 4061
Oy 3721 AGCACAAGT 3780
Db 4062 AGCACAAGT 4121
Oy 3781 TCAGT 3840
Db 4122 TGT 4181
Oy 3841 AACATTAGACTGAGTCAAGACTGTGACGACCGGGCGGCCCATCACTACTACATAT 3900
Db 4182 AATATCAGAACTGAGT 4241
Oy 3901 GGCAAATTCCTGTCCGATGAGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3960
Db 4242 GGCAGATTATCGCAATGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4301
Oy 3961 GAATGTCAAGCGGT 4020
Db 4302 GAATGTCAATGT 4361
Oy 4021 GAGACAGTGGGTGTGAGTCACTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4080
Db 4362 GAGACCGAGGCTCAGGCTAGT 4421
Oy 4081 ACCCCCAACCCCAATGAGAGGT 4140
Db 4422 ACTCCCAAGTAACTAGAGAGGT 4481
Oy 4141 GGGAGGGCAATTCCTGT 4200
Db 4482 GGGAAAGTATTCCTGT 4541

4201 AAGAAAAGTGTGACGAGCTGCGCGGCGCTTGGGGGTATGGGCTTGAATCAGTGGCA 4260
4542 AAGAAAGAGTCGACGAGCTGCGAGCGGCTTCCGGGGCAVGGGTGCAATGCGCTTGCA 4601
4261 TACTACAGAGGGTGAAGTCTCCGTAAATACAACTCAGGGAGAGTATGGTGTGCGCC 4320
4602 TACTATAGGGGTTCGACGTCTCCGTTATACCACTCAGAGAGAGCGTGGTGTGCGCC 4661
4321 ACCGAGCCCTCATGACAGGGTATACGTGGGACTTTGACTCCGTGATCGACTGCAAGTA 4380
4662 ACGATGCCCTTAAGACTGGGTACACGGCGACTTTAATCTGTCATGACGTGAATGTT 4721
4381 GCGGTCACTCAAGTTGTAGACTTCAAGTTAGACCCCAATTCACCATATCACAAGATT 4440
4722 GCACTCTCATGATTTGTGACTTTAGCCTTAGACCCAACTTCACCATACCACTCAAAAC 4781
4441 GTCCCTCAGAGCGCTGTCTCAAGTACGAGCGCGGGGTGCGACGGGTAGGGGAAAGCTG 4500
4782 GTCCCTCAGAGCGCTGTCTCCCGTATGCAAGTGAAGGAGAACTGGGAGGGGGCGATTG 4841
4501 GGCATTTATAGATATGTTTTCATCTGTGAGCGAGCTCAGGAATGTTTGAAGTGTGTG 4560
4842 GGGATTTTACAGATATGTTTGTGAGGAGGAGGCGCTGCGGATGTTGCAAGCGTAGTG 4901
4561 CTCTGAGTGTCTACGACGAGGGCGCGATGTATGAGCTCAACCATGAGAGACACC 4620
4902 CTCTGAGTGTCTATGATGTCGAGGGAGCGCTGTATGACAGCTTACACCTGCTGAACCTA 4961
4621 GTGAGGCTCAGAGGCGTATTTCAACACGCGCGTTTGTCTGTGTCMAAGACATCTTGAG 4680
4962 GTGAGACTCGGGGCTATTTCAACAGCGCGGGTTTGCCTGATGTCAAGACCATCTGGAG 5021
4681 TTTTGGAGGACGTTTTCACCGGCTTCAACACATATAGTCCCATCTTCTTCCAAACA 4740
5022 TTTGGGAAACGGCTTTTACAGGCTCACACATTTAACGCCCATCTTCTCCCAAGC 5081
4741 AAGCAATCGGGGAAAAATTTGCACTTTAAACAGCCCTACAGAGTCAAGTGTGCGTAGG 4800
5082 AAGCAAGAGAGAGAAATTTGCTATCTAAACGCGCTTACAGGCAAGGCAAGATGCGCAGG 5141
4801 GCCAAAGCCCCCGCGCTGTGGAGCGTATGTGAAGTGTGTAATCTGCACTCAAGCCC 4860
5142 GCAAAGGCCCTCTCTCTGTGGAGGATGTGGAAGTGTCTAACTAAGCTCAAACT 5201
4861 ACACTGTGGGCCCCACACCTCTCTGTATACCGCTTGGGCTGTGTACCAAGAGTCAAC 4920
5202 ACACTGACTGGTCCACCCCTCTCTGTATACCGCTTGGGTCGCGTGAACCAATGAGTAAAC 5261
4921 CTGACACATCCCGTGAAGAAATATCATGCGACCTGACATGCAAGCGACCTTTGAGGTATG 4980
5262 TTGAACGACCCCGTGAAGAAATATCATGCGACCTGACATGCAAGCTGACCTTGAAGATCATG 5321
4981 ACCAGCAGATGGGTCTTTGGAGGGAGAGTCTTGGCGGCGCTGCGCGGTATTTGCTTGGCG 5040
5322 ACAAGCTCATGGGTCTGCGGGGGGGGTCTAGCGCGGTGAGAGCTTACTGCTGGCG 5381
5041 ACCGGGTGTCTTTCATCATGCGCGCTTGCACATTTAAACAGGAGCGCTGTGGCGCG 5100
5382 ACTGGCTGCAATTTTCATCATGCGCGCTTACCTGATATGATCGGGGTGTGTGTRCCCT 5441
5101 GACAGAGAGGTCTCTATGAGGCTTTTGAAGATGAGAGGAATGTGCTTATGGGCGGCT 5160
5442 GACAGAGATCTTATATGAGGCTTTGATGAGATGAGAAATGCGCTTCAAAAGCGCGC 5501
5161 CTGATTTGAAGAGGGGAGCGAGATAGCCGAGATGCTGAAGTCCAAATCCAAAGCTTATTTG 5220
5502 CTGATTTGAAGAGGGGAGCGAGATAGCGAGATGCTGAATCTAAGATACAAAGGCTCTCTA 5561
5221 CAGCAAGCTTCCAAAGACATACAGACATACAAACCATGTGAGGCTTATGAGCCCAAG 5280
5562 CAAAGGCGCAAGGAGCTCAAGRCATRCAGCCAGACTATACAGTATCATATGAGCCCAAG 5621

5281 GTAGAACAATTTCTGGGCAAAACATATGTGAACTTATGCGGATCCAAATACCTGCA 5340
5622 CTGGAACAATTTTGGGCAAAACATATGTGAACTTATGAGGTATACAGTACTTACCA 5681
5341 GGAATATCAACACTGCGAGGAAACCTGCAATAGCTTCAATGATGGGCTTCAATGCCGC 5400
5682 GGAATCTCAACCTTACCGGAAATTCGCAATGATCAATGATGCTTTTATGAGCGCGC 5741
5401 CTGACAGTCCGCTGTCAACAACACCACTATCTCTCAACATTTTGGGGGCTGCTA 5460
5742 CTGACTAGCCCACTTACCAACACACCACTATCTCTTGAACATCATGAGGAGATGCTTG 5801
5461 GCATCCCAATTTGACACACCGCGGGGCGCACTGAGCTTCTGATGCTAGTGGCTATGAGGA 5520
5802 GCTCTATAGATGCGCGCGCTGCGAGGACCACTGAGCTTCTGATGCTATGATGAGG 5861
5521 GCTGCGGTAGGAGTATAGGCTTATAGTATGATGATGATGATGATGATGATGATGATGAT 5580
5862 GCGGCGGTGCGAAGCATAGGCTGAGGTATGATGATGATGATGATGATGATGATGATGATG 5921
5581 GCGGCGATTTTGGGGGCTCTGCTGCAATGATGATGATGATGATGATGATGATGATGATG 5640
5922 GAGGCAATTTCAAGGGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATG 5981
5641 GAGGATGCTGCAACTTGTGCTGCTGCAATGATGATGATGATGATGATGATGATGATGATG 5700
5982 GAGGAGCTTGTGATATCTCTGCTGCTGATATGATGATGATGATGATGATGATGATGATG 6041
5701 ATCTGCGCGGCAATTTCTGCGCGCAACGTGGGAGCGGGGAGAGCGCGCTTCAATGATG 5760
6042 ATCTGTCAGACATATGCTGCGCGCGCAAGTGGTCAAGGAGAGGAGGAGGAGGAGGAGG 6101
5761 AATGACTATTTGCTTGTGCTTCAAGAGAAATCAAGTGGCGCGCGCGCGCGCGCGCGCG 5820
6102 AACAGACTGATGCTTGTGCTTCAAGGAGAAACAGTGGCGCGCGCGCGCGCGCGCGCG 6161
5821 GAGTGGATGCTGTCGAGGAGGTGATGACCACTACTTGGGCTTCTTACCATATGAGGCTG 5880
6162 GAGTCTGACGCTTCAACAGGCTGTACGAGGCTCTGAGTTCACTTACATTTACAGCTTA 6221
5881 CTGAGAACTCCCAACTGATTTACTGAGGAGCTGCGCGCATTCATTCAGCGCGGCTG 5940
6222 CTGAGGACTATGCTGCGGATGATCACTGAAGATGCGCAATTCATTCAGCGCGGCTG 6281
5941 CTGCGGATGTGTGGGAGCTGCGGTTTGCACCATCTTCAACAGACTTTTAAATTTGGCTG 6300
6282 CTCAGAGCAATTTGGGATTTGGTTTGTTCATCTTCAACAGACTTAAATTTGGCTG 6341
6001 TCCAAATTTATTCCAAGATGCGCGGCTTCTCTGCTCTCTGCTCTCTCTCTCTCTCTCT 6400
6342 TCAGAAATTTACTCCCAAGATGCGCGGATTCCTTTATCTCTTCCAGAGGAGATACAG 6401
6061 GCGGTGTGGCGCGGCACTGCGATCATGACCAACAGGTGCTTGTGCGCGCGCAATATCTCT 6120
6402 GGTGTATGGGCTGTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 6461
6121 GGCATATGCTGCTTGGGCTTCAATGAGATCAAGGAGCTTAAAGCTGATGATATCTG 6180
6462 GGCATATGCTGCTTGGGCTTCAATGAGATCAAGGAGCTTAAAGCTGATGATATCTG 6521
6181 CAGGGAGCTTTCTTCAATTTTACAGGAGGCGAGGTGCTGCGGAGAAACCGCGCGCA 6240
6522 CAGGGAGCTTTCTTCAATTTTACAGGAGGCGAGGTGCTGCGGAGAAACCGCGCGCA 6281
6241 AACTTTAAGGTGCACTGAGAGGTGCGGCTTCAAGTATGCGGAGTATGAGTATGAGTATG 6300
6582 AATTTAAGAGCGCAATTTGAGAGGTGCGGAGTATGAGTATGAGTATGAGTATGAGTATG 6441
6301 GGTATATACATCACTTAAAGAGCTTCAAGTATGAGTATGAGTATGAGTATGAGTATG 6360
6442 GGTATATGCTGATATGATGAGGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 6401
6361 CCTCTCCGAGATTTCTTTCTGAGTGAAGAGTGCAGATCCATAGTGTGCGCGCGCA 6420

Db 6702 CCAGCTCAGAAATTTTCTCTTGGGTGGACGGGGGGAATTCACCGATTCGCCCCGTW 6761
QY 6421 CCGAAGCCGTTTTTCCGGGATGAGGTCTCGTTTCGCGTTAGGGCTTAATCATTTGTGTGC 6480
Db 6762 CCAAGTCCCTTCTTCCGGGATGAGGTAACTTCACCGTAAAGCCCTTAATCTCTTGTTGATC 6821
QY 6481 GGGTCCGAGTTCTTGGGACCTGAAACCGACACAGAGTATGATGATCCATGCTAACA 6540
Db 6822 GGCTCTAGCTCCCTTGGATCTGTAGCGGACACCGAGTACTGGCTCTCATGTTTACA 6881
QY 6541 GATCATTTCAATACCGGGGAGACTGACAGCGGGCTTTAGCGCGGGGTCAACCCCA 6600
Db 6882 GACCCGTCCACATCACGCKGAGCGGACCGACGAGCGATTGGCAAGGGGATCTCCCCY 6941
QY 6601 TCCGAGGCAAGCTCCTGGGCGAGCGAGTATCGGACCATGCTGCGAGCCAGCTGAGCC 6660
Db 6942 TCACAGGCTAGCTCTCAGCGAGCCAGCTCTGTGCCCTCTTGAAGGCTACTGTATCC 7001
QY 6661 ACCCAAGCAAGCCTATGATGTGACATGTGTGATGCTAACTGTTCATGTGGGGGCGAT 6720
Db 7002 ACCCATTAAGACAGGATATGATTTGTGACATGTGTGATGCTAACTTTTCATGAGAGGAT 7061
QY 6721 GTGACTCGGATAGAGTCTGGGTCCAAAGTGGTCTGTGACTCTTCGACCTTCGACATGATC 6780
Db 7062 GTGATCCGAGATTGAGTCTGACTTAAGGTGATCGTTCTAGACTCCCTGATTCATGACT 7121
QY 6781 GAGAGAAAGAGGAGCCTTGAAGCTTCGATACCATGAGATACATGATGCTCCCAAGAGAGG 6840
Db 7122 GAGGTAGAGATGATCTGAGCTTCTGTACCATGAGTACCTGATACAGAGAGAGAG 7181
QY 6841 TTCCACAGGCTTTAACCGGCTTGAGCAAGGCTGATTAACAACCGACTTGTGAAATCG 6900
Db 7182 TTCCACAGGCGGCTCTCTTGGGCCGCTGCAAGCTAATCCGTTTGTATGAGAGCA 7241
QY 6901 TGGAAAAGGCCAGATTACCAACCGGCACTGTGGGGCTGTGCTCTCTCTCTAGG 6960
Db 7242 TGGAGAGGCGGGCTATGAAACCAACCTGCTGAGCTGTGCCCTCCGCCCAACCT 7301
QY 6961 AAAACCCGAGCGCTCCGCCAAGAGAGGAGGAGCAGTGGGCTTAAGTGAAGATCCCATTA 7020
Db 7302 CAAGGCAAGTCTCTCACTCGAGGCGCGCTGAATATCTGACCCAGAGACATATGTG 7361
QY 7021 GAGAGTGCCTTCAACAGCTGCCCATTAATCTTGTGCGAGCCGCCCAAGCGGCGAT 7080
Db 7362 GAGGGGRTCTTCAGGAGATGAGTGAACAAGTGTCAAGCCCTCCCAAGACAAATGAC 7421
QY 7081 TCAGGCTTTTCAAGGGGCGGGCGCTGCCGATTCGGGAGTCAAGACGCTCTGATGAG 7140
Db 7422 TCCGCTCACTCACTGAGCGGATACCGAGAGAGCATGCTCAAGCAACCTCTGACGAG 7481
QY 7141 TTGGCCCTTTCGAGAGAGGTTCCATCTCTTCATGAGCCGCCCTGAGAGGGAGAGCTTGA 7200
Db 7482 ACTGCGCTTTCAGAGCGGGGCTCACTGTCTCTCACTCTCTCCCTTGAAGAGAGCGGGA 7541
QY 7201 GATCCAGACTGAGGCTGAGCAGGTAGAGCCCAACCCGCCCAAGGGGGGGTGGCA 7260
Db 7542 GACCCYGACCTGAGATTGAAACAGTGGGATCCGGTCCCTTGAAGGGGAGTGTGAG 7601
QY 7261 GCTCCCGGCTGGGATCTGGGGTCTGTGTCTACTTGTCTCGAGAGAGAGATCTCGTGTG 7320
Db 7602 GTCATTTGATTCGAGCTCTTAAGTCTGTGTCTCAAGTCTCTGATCAAGAGAGATTCTGTATTC 7661
QY 7321 TGCTGCTCATGTCTATCTCTGAGCGGGGCTCAATATACCTCTGTAAGTCCGAGAGAG 7380
Db 7662 TGCTGCTTAATGTCTATCTCTGAGCGGGGCTCTCATTAACAATGTGGGCGCGAAGAG 7721
QY 7381 GAGAGATTACGATTAAACCCCTTGAGCAATCCCTGTGCGATATCAACAAGGTGAC 7440
Db 7722 GAGAGATTACGATTAAACCCCTGTAGTAAATTCGCTCATGGGTTCCATAAAGGTGAC 7781
QY 7441 TGTACCAACAAGAGGCGCTCACTAAAGGCTAAAAAGTACTTTGTGATGATGCA 7500

Db 7782 TCCACAACCTCGAGAGTGCCTCTCTGAGGGCAAGAGAGTACTTTTGACAGGGTGCAG 7841
QY 7501 GTGCTGAGCTCTCACTAGAGCTCAAGCTTAAGAGACATTAAAGCTAGCGGCTCAAGGTC 7560
Db 7842 GTGCTGAGCGACACACTATGACTGACTTGTGAGAGGTTAAAGGGGCGCTCTTAAGTGT 7901
QY 7561 ACCGCAAGGCTCCTCACTAGAGAGGCTTGCAGTTAAGCCCAACCCATTCTGCAAGA 7620
Db 7902 RGTGCAAGGCTCCTCACTAGAGAGGCTGCGGCTGACCCCGGCCACTCCGCAAA 7961
QY 7621 TCTAATTAAGGTTGGGGCTAAGAGGTCCGCAAGTTCTCCGGAGGGCGCTTAACAC 7680
Db 7962 TCGCATAGGATTGGGGCAAGAGGTGCGCAGCTTAATCAAGAGGGCGCTTAACAC 8021
QY 7681 ATCAATCCGTTGAGAGAGACCTCTGAGAGACTCAGAAACCAATTCACCAACAT 7740
Db 8022 ATCCGGTCCGTGTGGAGAGACCTCTCGAAGACCAACATTCACCAACTATC 8081
QY 7741 ATGGCAAAAATGAGGTCTTGTGCTGACCCCAACCAAGGGGGCAAGAAAGCTGCG 7800
Db 8082 ATGGCTAAAAATGAGGTCTTGTGATTCACATTAAGGTGGGAAAAAGCAGCTCCG 8141
QY 7801 CTATGCTTACCTGAGCTCCGCGCTCAGAGTCTGCGAAGATGAGCCCTTATGACAT 7860
Db 8142 CTGATGTATACCCCGACCTTGGGTCAGGGTGTGAAAGATGGCCCTCTATGACATC 8201
QY 7861 ACACAAAACTCTCAGAGCGGATGAGGGGCTTCTTATGATTCAGATTTCCCGCT 7920
Db 8202 RCACAAAACTTCCCAAGCATTAATGGGCGCATCTTAAGGTTCCAAATCTCTCCGCA 8261
QY 7921 CAGCGGTAGATTCTTTGAAAAGATGGCGGAAAAAGAGACCTTAATGGGTTTTG 7980
Db 8262 GAACGGGTGATTTCTCTCCCAAGCTTGGGAAATGAAGAACCAATGGGGTCTCG 8321
QY 7981 TATGATACCGAGTCTTGAATCAACCTCACTGAGAGAGACATCGAGCTAGAGGCTC 8040
Db 8322 TATGACACCGGCTTGTGATCAACCTGACAGAGAGAGACATTAAGAACAGAAATCC 8381
QY 8041 ATATATCGGGCTGCTCTTGGCCGAGAGGCGCCACATGCACTAGCTAGCTAG 8100
Db 8382 ATATATCAGGCTTGTCTGCTCTCAAGAGCCAGAACCTGTATATGTGTGTGAGAC 8441
QY 8101 AGACTTACGTTGAGAGGCGCTATGTTCAACAGCAAGGCCCAACCTGCGGTTACAGGCT 8160
Db 8442 AGACTTACGTTGAGAGGCGCTATGTTCAACAGCAAGGCCCAACCTGCGGTTACAGGCT 8501
QY 8161 TGGCGGCAAGCGGGTGTCTACCACTAGAGAGGAGAACCATCATGCTACGTAGAA 8220
Db 8502 TGGCGGCAAGCGGGTGTCTACCACTAGAGAGGAGAACCATCATGCTACGTAGAA 8561
QY 8221 GCTTAGCGGCTTGTAAAGCTGAGGAGTAAATGCGGCCCAATGCTGTATGCGGCGAT 8280
Db 8562 GCCCTTTCAGAGCGTGTAAAGCTGCGGAGTGTGAGACCTGTATATGTGTGTGAGAC 8621
QY 8281 GACTGTGTCTATCTCAAGAACCCAGGGAGCCGAGAGAGAGAGCGGAACTGAGAGCC 8340
Db 8622 GACTGTGTCTATCTCAAGAACCCAGGGAGTAAAGAGAGAGAGAGCGGAACTGAGAGCC 8681
QY 8341 TTCAAGGAGGCTATGACAGAGTATTCGCGCCCTCTGTGAGACCCCGCCAGACCGGAT 8400
Db 8682 TTCAAGGAGGCTATGACAGAGTATTCGCGCCCTCTGTGAGACCCCGCCAGACCGGAT 8741
QY 8401 GATCTGAGGCTGATTAATCATTTGCTCTCAATATGTCTGTGAGCGCTGAGCCCAAGG 8460
Db 8742 GACTGTGAGGCTTAAATCATCTGCTCTCAAAAGTATCGGTAGCGCTGAGCTCTCGGGT 8801
QY 8461 CGCGCAGATTAATCTGACAGAGACCTTACCACTCAATGCGCGCGGCTGCTGGAA 8520
Db 8802 CGCGCAGGATTAATCTGACAGAGACCTTACCACTCAATGCGCGGCTGCTGGAA 8861
QY 8521 ACAATTAGACATCCCTCTGCAATTCATGCTGTGGGAAACATCATCCAGTACGCGCGGAC 8580
Db 8862 ACAGTAAAGACATCCCTCTGCAATTCATGCTGTGGGAAACATCATCCAGTACGCGCGGAC 8921

```

OY 8581 ATATGGGCTCCGATGGTCTGTATGACAACATTCTTCTCATTTTCAATGGCTCAAGACAG 864
Db 8922 ATCTGGGCTCCGATGGTCTGTATGACTACTTCTTCTCATATCTATTTGGCCAGACACT 8981
OY 8641 CTGACACGAACCTTCACATTTTGAGATGTACAGGAGCGGTGTACTCTCGTAGTCCCTTGGAC 8700
Db 8982 CTGAACCAAAATCTCAATTTTGAGATGTACGGGCGAGTATCTGGGTCAATCTCAATTTAGAC 9041
OY 8701 CTCCAGCTATATTTGAAAGGTACATGGGCTTGAACGTTTTTCTCTGACACATCACT 8766
Db 9042 CTACCGGCAATATTTGAAAGGCTACATGGGCTTGAAAGCTTTTTCATGACACATCACTCT 9101
OY 8761 CCCCACGAACGTGACACAGGGTGGCTTACAGCCCTCAGAAAACTTGGGGGCGCACCCCTCAGA 8820
Db 9102 CCCCACGAACCTCTACCGGGTGGCAGCACTCTCAGAAAACTTGGAGCGGCTCTCCCTTAGA 9161
OY 8821 GCGTGAAGAGCCCGGACAGTGCAGTACAGGCGTCCCTCATCTCCCGTGGGGGAGAGCG 8880
Db 9162 GCGTGAAGAGTCCGGGCGCGTGCAGGCTTGAAGCTTCACTCATCTGCGCAAGAGAGCGAGCG 9221
OY 8881 GCCGTTTGGCGGTCCATATCTCTTCAATTGGGCGGTGAAGACCAAGCTAAACTCACTCA 8940
Db 9222 GCCATTGTGTGCGCGGTACCTCTTCAACTGGGCGGTGAAGAAACAAAGCTCAACTCACTCA 9281
OY 8941 TTGCGCGAAGGCGCGCTCTGTGATTTATTCAGCTGTTCACGTCGGGCGCGGCGGGGCG 9000
Db 9282 TTGCGCGAAGGCGCGCTCTGTGATTTATTCGGGTGTTCACGTCGGGCGCGGCGGGGCG 9341
OY 9001 GACATTTATCAACAGCGTGTGCGGTGCGCCGACCCGCTATTGCTCTTTGGCGCTACTCTA 9060
Db 9342 GACATTTATCAACAGCGTGTGCGAGTCGACGCCGCTATTACTCTTTGGCTACTCTA 9401
OY 9061 CTTTGTGTAGGGGTAGGCTTTTCTACTCTCCCGCTCGG 9099
Db 9402 CTTAGCGTAGAGTAGGCACTCTTTTACTCTCCCGCTCGG 9440

RESULT 4
US-07-925-695-6
; Sequence 6, Application US/07925695
; Patent No. 5428145
; GENERAL INFORMATION:
; APPLICANT: OKAMOTO, Hiroaki
; APPLICANT: NAKAMURA, Tetsuo
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOME,
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beveridge, DeGrandi, Weillacher & Young
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: US
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/925.695
; FILING DATE: 19920807
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 287402/91
; FILING DATE: 09-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 360441/91
; FILING DATE: 05-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Weillacher, Robert G.

```

/		REGISTRATION NUMBER: 20,531	
/		REFERENCE/DOCKET NUMBER: 06/87-48009	
/		TELECOMMUNICATION INFORMATION:	
/		TELEPHONE: (202) 659-2811	
/		TELEFAX: (202) 659-1462	
/		TELEX: WUI 64470	
/		INFORMATION FOR SEQ ID NO: 6:	
/		SEQUENCE CHARACTERISTICS:	
/		LENGTH: 9511 base pairs	
/		TYPE: NUCLEIC ACID	
/		STRANDEDNESS: single	
/		TOPOLOGY: linear	
/		US-07-925-695-6	
Query Match	62.1%;	Score 5651;	DB 1; Length 9511;
Best Local Similarity	59.4%;	Pred. No. 0;	
Matches 5403;	Conservative 1567;	Mismatches 2129;	Indels 0; Gaps 0;
OY	1	ATGAGCACAAATCTTAAACCTCAAGAAGAAAAACAAGAAAACCACCAACCGTCGCCACAA	60
Db	342	AUGAGCACAUAUCCUAAAACUCAAGAAAAAACAAGAAAACAACAAACCGCGCCACAG	401
OY	61	GACGTTAAGTTTCCGGGCGGGGCAGATCGTTGGCCGAAGATACTTGTCGCCGCAAG	120
Db	402	GACGUUAAAGUCCCCGGGGGCGGUCAGAUCCUUUGCCGAUUUACUUGUCUCCGCGAAG	461
OY	121	GGCCCCAAGTTTGGGTGTGTCGCGCACAAAGAAAGACTTCAGAGACGGTCCACAGTGA	180
Db	462	GGCCCCAAGUUGGUGUGUGCGGCCAACAAAGAAAGACUUCYGAAGCAUCCAGCCGCUUGA	521
OY	181	AGCGCGCAAGCCCATCCTAAAGATCGCGCTCCACTGGCAAAATCTTGGGGAAAAACAGGA	240
Db	522	CGAGCGCAAGCCCAUCCGAAAGAUCCGCGCTCCACCGCAAGUCCUGGGGAAAAACAGGA	581
OY	241	TACCCCTGAGCCCCTATACGGGAATGAGAGGAATCGAGCTGGGAGGATGGCTCTGTCGCC	300
Db	582	UAUCCUUBGCCCCCUGUACGAAACGAGGUGUGCGGCUBGCGGUGGCTUCUGUCCCCC	641
OY	301	CGAGTTTCCGTCCTCTTGGGGGCCCAATTGACCCCGGCAATAGTCGCGCAACGTGGGT	360
Db	642	CGCGGUCUCGUCUACUUGGGGCCCAACCGACCGACCCCGGCAUAAUAVACAAGCAUUVUGGC	701
OY	361	AAGTCATTCGATACCCTTAACGTGGGCGCTTTGGCCGACCTCAATGGGGTTCATCCCTGTG	420
Db	702	AGAUUUAUCCGAUACCAUUAACGUUGUGUUUUCCCAUCUACAGGGGUAUACUCCUGUGUU	761
OY	421	GGCGCCCCGCTCGCGCGCGTGCACAGACTTCGCGCATGGCTGAGAGTCTTGAAGAC	480
Db	762	GGCGCCCCGGUYGGAGGGCGUGCCAGAGCUCUUGGCAACAGGUGUAAGGUCUCUGGAAGAC	821
OY	481	GGGTTAATTTTGCACAGGGAATTACCCGGTTGCTCTTTTCTATCTTCTTGCTGGCC	540
Db	822	GGGUAUAUUDCGMAACAGGGAAUUAUACCCGGUUGUCUUUUUUAUCCUUUUUUCUGUCU	881
OY	541	CTGCTGTCGTCGATCAACACCCCGGCTCCGGCTCGAGTCGGAAGTGAATCATGATCCGCG	600
Db	882	CUUCUGUUAUCGUCACARUCCACAGUGUCUGACAGGAAAGUCAGAAACAUVAUUGUUAAGC	941
OY	601	TACATGATGATTAACAATGACATGACCAATGACAGATTACTTGGCAGCTTCAGGCTGCTC	660
Db	942	UACUACGCCACUAUAVUUGUCUCAAACAACAGCAUACUCUGGAGCAGUCAUGAGCGCAGUU	1001
OY	661	CTCCACGTCCTCCGGGTGCTCCCTGCGAGAAAGTGGGGAAATGATCATCTCACTGCTGATA	720
Db	1002	CUCCAUUCUUCUGGAUUCGUCCCAUUGUGAAGAAVAAUAYAGCAACUUGCRUUGCUUGAUA	1061
OY	721	CCGGTCTCACCGAATGTCGCGGTGCAACGCGCCCGGCGCCCTCAACGACAGGCGTTGCGGACG	780
Db	1062	CAAUAAACACCCRRACGUGGCUUGAUAACCCCGGUGGCGUCAUCUGAAGCUCUGCMAACA	1121
OY	781	CACATCGACATGGTTTGATGTCCGCAACGCTGTGCTCTGCGCTCTACATGTCGAGGAGACCTC	840
Db	1122	CACUCGACAUUGAUCGUAUGGACGUAACGCGCTGCGCTGCGCTGUGUUGUGGAGAUUG	1181

OY	841	TCGCGTGGGGGAGTGCCTGACGCCCCAAATGTTCACTGCTCGCGGACACCACTGGTTT	900
Db	1182	UCCGGGGCCGGBAUBAUYCUAUCGACGGCCUUBAUGUAUACCAACAAGCCCAACAUCU	12411
OY	901	GTCCAAAGACTGCAATTGCTCTCATCTTACCTGGTATCATCTGTGACACCCGATGGACTG	960
Db	1242	ACCCAAAGAGUCACAUCUUCACUACCAAGUUCACUACACCGGCAUUGCAUGGCAUUG	1301
OY	961	GACATGATGATGAACCTGGTCCGCCACGGCTACCATGATCTTGGCGTATCGCATGGTCTC	1020
Db	1302	GACAAAGAGCCRRARCUGGUCUCCCAACUUCURCCAGAACCUCCGCUAAGCVCUCUGYU	1361
OY	1021	CCCGAGGTCAATTATAGACATCTTATGGGGGCTCATTTGGGGGCTCATTTGGGCTTTGGC	1080
Db	1362	CCCCARCUGGUCUCGAAAUUUAUYUCGGCGGCAUUGGGUGUGYUUYUGCUUGGSC	1421
OY	1081	TACTTCTTATGACAGAGAGCGTGGCGGAAAGTCGTTGTTCATCTTCTGTTGGCGCGCGG	1140
Db	1422	UAUUUCUCACUAGACAGAGCGUGGGCCAAAGUCRUYGCAUUCUCUCUCUUGUUGCGGA	1481
OY	1141	GTGGAACGGCGCACCCATATCTGTTGGGGGTTCTGCGCGCAGACACACGGGCGCTCACCC	1200
Db	1482	GUGGAGUGCAACCAUUCUCCACGCGYAGSAAAGGGGUCURCCGCKMKMGGGMUCUCU	1541
OY	1201	AGCTTATTTGACATGGGGCCCCAGGAGAAAAATCCAGCTCGTTTAAACCAATATGGACGTG	1260
Db	1542	RGCCUCUUAUAUCUGUGGCCAAGACGAACUCUUAUURAUACACCAUUGCAGCUGG	1601
OY	1261	CACATCAACGCGACCGGCTTGAACTGCAATGAATCTTCCTTGCAACCGGCTTATCGCGTCT	1320
Db	1602	CACAAUAAACCGGACUGCCCTCUCAUUGCAUAGACAGCUIASAGAAGGCUUUCUGCUCUC	1661
OY	1321	CTGTTCTTACCCCAAGCTTCAACTCGACAGATATGTCGGAAGTATCCGCGCTGCGCGC	1380
Db	1662	YUGKUUUUCMCCRCARUUCAACGUCUGCGGCGGCCCGAGCGCUCUUCUGCGCGC	1721
OY	1381	AGTATCGAGGCGCTTCCGGGGTGGGATGGGGCGCTTGCAATATGAGAGATATGTCAACAT	1440
Db	1722	GGGCUUGACGAYUUYCGCAUCGGCUGGGGAAACUUGAUAUACGAACCAAGUCACCAAC	1781
OY	1441	CCAGAGATATGAGACCCCTATTTGCTGGGACTAACCCACCAAGGACATGTGGCGTGTCTC	1500
Db	1782	GAUGGACCAUGAGGGCCGUACUGCUGGCAUUAACCCCGAGGCGUUGGGCAUUGUCGG	1841
OY	1501	GCGAAGACTGTGTGTGGCCCAATGACTGTTTCAACCCCAAGCTGATGTAGTGGGACAG	1560
Db	1842	GCUAAGACGUGUUGGGAGCCGGUCUUAUUGYUACCCCUAGCCUGUUGUGUGGGCACCC	1901
OY	1561	ACCGACAGGCTTGGAGCGCCCACTTACACGTTGGGGGAGAAATGAGACAGATGTCTTCTTA	1620
Db	1902	ACUGACAAAGCAGGGGCGUACCCACCUACACUUGGGRRGAAACGAGACCGAUUCUUCUG	1961
OY	1621	TTGAACAGACTCGACCAACCGGTGGGTCATGTGTGCGCTGCACGTGATGAATCTTTCT	1680
Db	1962	CRAAUUACACAAGACCCCGGAGAGGACUUGUUGCGCTUGACAYUGAUUGAACGGAGCU	2021
OY	1681	GGCTACACCAAGACTTGGGCGGACACACCTGCGGCTATAGAGCTGACTTGAACCCGACG	1740
Db	2022	GGGUUUCAUAAACAUUGCGGUCACACCUUGCCGCAUUDAGAAAGAUACAACAGCACU	2081
OY	1741	ACGGAAGCTGTTGTGCCCCACGGAAGCTGTTTAGAAGACATCTGATACCACTTACCTCAAA	1800
Db	2082	CUCGUAUUUAUUGGCCCCCAACACUUGUUUAGAAAGCAACCAAGAUUCUAUCUUAUAG	2141
OY	1801	TGCGGCTCTGGGCGCTGCTACGCGCCAAAGTGCCTGATGACTACCCCTACAGGCTCTGG	1860
Db	2142	UGUGGAGACAGGCGCUUGGUUAUCUCCAGUUGCCUGUAGACUAAACCUUAUAGRYUGG	2201
OY	1861	CATTACCCCTGACAGTTTAACTTAACTTCAATCTTCAAAATAAGATGATATGTGGAGAGGTT	1920
Db	2202	CAUUUUCGUGGACUGUAAACUUCACCAUCUUYAAGCGCGGAGUUAUGUAGAGAGGUG	2261

OY	192	GAGACAGGGCTCAAGGGCTGATGGAATTTCACTGTTGGGATCGTTTCAACTTGGAGAC	1980
Db	2262	GAGCAUGAGUUTCUGCAGCAUGCAUUCACGGCGGGAUUGCUCAGACUGGAAGAU	2321
OY	1981	AGAGACAGAAATCACTGTCTCTTTGGTGCATCCACACCGAATGGGCATTTTACT	2040
Db	2322	AGGGAUAGGGGVCAGCAGAGUCCACUCGUCUACUUCACUACUGAGUGGGCGUUYUCCA	2381
OY	2041	TGCTCTTAATCCGGAACCTGGCCGCTTGTGCATGTCCTTCCACCTCCACCAAAATC	2100
Db	2382	UGCCUCUCUCUGACCUACAGACUWCCACUGGCUAUNUGACUCCACCAAAACUUC	2441
OY	2101	GTGACGCTACAAATTCATGTATGCTTATGCTTCACTTGCCCTCCCAAAATACATCTGTCGATG	2160
Db	2442	GUAGACUGGAGUACUUYACGACUUYUCUGGCUUGACAAAGUACUUCUGUAGUUG	2501
OY	2161	GAGTGGGTAAATCTCTTAATTCCTGCTTTAGCGGACGCGAGGGTTTGCGCTGTTATG	2220
Db	2502	GAGGGGUGAUCUCCUCUUYUCUUYUUGGCAACCCAGGRUCUUGGACUUGCUCUUG	2561
OY	2221	ATGCTCATCTTGTGGGCGAGGCGGAGACGACATGAGAAGCGGTCACTTTCAGACT	2280
Db	2562	AUGUCUACMAUACUGGGCCAGGCCGAGCGCGUAGAAAGCUACUUCUUGACUCC	2621
OY	2281	GGAGCGCAGCTAGCTGSCAATGCTTCTTAATTTGTCACTTTTTCGTGGCTGTTGG	2340
Db	2622	GCUGAYCUGUGAUGGCCAAUUGGUCGCCUGUGUUYUACUUCUCUUAACGCGCCUGG	2681
OY	2341	TACATCAAGGGTGGGGTAGTCCCTTTAGTCACTTATCCCTCACTGGCTGTGTCTTT	2400
Db	2682	UACUAAAAGGCGAGGUGUGUCCCGUGGCGACGUACUCUGBCUGGCUURUGUCCUC	2741
OY	2401	AGCCTACTGCTCTTAGCATTTGCCCAACAGGCTTATGCTTATGACGATCTGTGCATGCG	2460
Db	2742	CUCCUCUUAUCUGGCGUYUACACACAGGCUUAVGCCUUGACGCGUCUGAACAAAGG	2801
OY	2461	CAGATAGAGGGGCTGCTGTGTAAATGATCACTCTTTTACTCTACACCCCGGTTAAAG	2520
Db	2802	GAACUGGGCGUGGCAUUAAGUAAUUAACCAUCUUAACUUCUACCCGACGACUACAG	2861
OY	2521	ACCCTTTCAGCGGTTTGTGTGTGTGTGTGCTATCTTTCGACCCCTGGGGAAGCTATG	2580
Db	2862	AUCCUCUGAGCGCUBUAGUUGUGUGCUCUACUAGCUGUUCUUGGCCGAGCCAG	2921
OY	2581	GTCCAGAGTGGGACCAACCTATGCAGGTGCGCGGTGGCGGTGATCATATATGGGC	2640
Db	2922	AUUCAGCAAUUGGUYUCCCCCUCUGAAGGUCGAAAGGGGGCGUGACGGGAUUCUGGUG	2981
OY	2641	GTGCCAATATCTTCAACCGAGGTGTGTGTTGACATTAACCAAGTGGCTCTTGGCGGTGCTT	2700
Db	2982	GCUGUACAUCCACCCACGCGCTUGUGUUBUAGGUCACGAAUAGUUGUUGAUAUCCUG	3041
OY	2701	GGGCGCTGTTACTCTTAAAAAGTCTTTTGAACGCGCGTGCCTTCTGTCAGGGCTCAC	2760
Db	3042	GGGCGUGCUACUCUCUUAARAGCUCUCUGACGAAUCCGUACUUYUGUAGGGCCAC	3101
OY	2761	GCTCTACTGAGGAATGTGACCAATGCAAGGATCTGCGGGGGGCAAGTACGTCACATG	2820
Db	3102	GCUUYUCUACAGUGUGUACCCUGAGUAAACACUCCRCRGGGUGUAGUAAUCUACAGUG	3161
OY	2821	GGCGTACTAGGCCCTTGGCAGAGTGACTGGCACTTACATCTATGACCACTCACCCCTATG	2880
Db	3162	CUGUURAUACACUAGGCAUUGAGACGGGCAUUAACUACUACGACACACUCUCCCUUA	3221
OY	2881	TGCGATTGGGCTGTAATGTGCTGCGGACCTGGCGGTGCGCGTTGAGCCTATCATCTTC	2940
Db	3222	UCAACUUGGGCGGCGCCAGGGUYUACGGGACCUUGGCAUUGCGCGUGAGCCUGUGUUYUC	3281
OY	2941	AGTCCGATGAGAAAGAAAGTCAATTGTCTGGGGAACGGACAGCTGCTGTGGGGAATTT	3000
Db	3282	AGCCCAUAGGAAAGAAAGUUCAUUUGUGGGGGGCGUGAAGUAGUGCGUUGGAGACUUC	3341
OY	3001	TTAACCGGACTCCCGTGTCCGCCGACTTGGTGGGAGGTCTCTTGGCGCCAGTGT	3060

Db 3342 CUGCAUGGCCUCGCGGUCUCGCGAGGCUAGUAGGAGUUCUGUGCGGCCGCGAC 3401
Qy 3061 GGCATATCTTCGAAGGGGTGAGATCTTTCGCCCCCACTCACTGTTACGCCAGACAGA 3120
Db 3402 GGGUACACCCGCAAGGGGUGGAAACCTUCUAGCCUUAUCUGUCUUAACUCAGCAAAACU 3461
Qy 3121 CGTGGCCTTTTGGGACCATAGTGTGATGATGACGGGGCGCGACAAGACAGAAAGACT 3180
Db 3462 CGUGUCUCUGGGUGUGUAGUGUGUAGCCUAGCGGGCGCGACAAAAGAGAGAGGCU 3521
Qy 3181 GGGGAAATTCAGGCTCTGTCAAGTCACTCACTCTCGGAACATCATCTCGGG 3240
Db 3522 GGGCAGGUCGAGGUCUGUCUGCGUCACAACAAUCUUCUGGGACAUUCAUUUGGGC 3581
Qy 3241 GTTTTGTGACTGTCTACATGAGAGCTGGACAAGACTGTGCGCGGCTACAGGGTCCG 3300
Db 3582 GUCUCUCGAGCAUUAUUAUCAGGGGCGUGUAUAAGACUUGCGCGGCCCAAGGACCA 3641
Qy 3301 GTCAAGCATGTACTCAGTGTGAGAGGGGACCTTAGTGGTGGCCAGCGCCCTGGG 3360
Db 3642 GUACUCAGUAGUACACAGCGACAGAGGGGACCUUGGGAGUAGCCUAGUCCCGCGG 3701
Qy 3361 ACTAAATCTTTGAGCGGTGACAGTGTGAGCGGTGACCTGTACCTGTACGCGAGAC 3420
Db 3702 ACUAGUCAUUGAGCCCGUACUCUGGGGCGGUAGACCUUAUCUGUACCCGAAAC 3761
Qy 3421 GCTGATGTCACTCCCGGTGCAAGACGCGGGGACAAAGCGGGAGCGCTACTCTCCCGAGA 3480
Db 3762 GCUAGUUCAUUUCGCGCCGAGAGAAAGUAGCCGAGGGGUGUAGCUUCUGCCCAAG 3821
Qy 3481 CCTCTTTCACCTTGAAGGGGCTCTCAGAGGCGCGGTGTATGCCCCAGGGGCACT 3540
Db 3822 CCCCUCUACACCUCAAGAGAUCAUCCGAGGGCCGUGUCUCUAGGAGGACACGCC 3881
Qy 3541 GTGGAATCTTCCGGGAGCTGTGTCTCTCGGGGCGTGTAACTCATAGATTTTCATC 3600
Db 3882 GUGGGCUUUGUAGAGGCGCGUUGUAGCAGGGGUGUAGCAAUUAUUGACUUCUAVC 3941
Qy 3601 CCGGTGAGACACTCCGACATGTCAAGCGGTCCCCCACTTTAGTGAACAAGACACCA 3660
Db 3942 CCGUGUAAUCACUCGACUURUCGCAACGCAAGCCCGGUGUUCUGUACACAGURCCCG 4001
Qy 3661 CCGTGTGCGCCGAGCTTATCAGTGTGGGTCTGTGATGCGCCGACGTGGCAGTGAAG 3720
Db 4002 CCGCUGUGCCCAAGUUCUACAGUGGGUUAUCUGACACCAAGCAGCAGGAGAG 4061
Qy 3721 AGCACAAGTTCCTGTGCGATATGTCTCAGGGGTATAAAGTGTAGTCTTAATCCC 3780
Db 4062 AGCACAAGGUCCTCGCGCGUUAUGCAUICAGGGGUAUAUAUUAUCUGUACUAAUCC 4121
Qy 3781 TCAGTGTGCGCACCTCGGGGTTTGGGGCGTACTGTCTAAGGACATGGCATCATCC 3840
Db 4122 UCUUGUGCGGCGCACUCUGUGUUGGGGCUCAUUGICCAAAAGCCCGAGUACACCU 4181
Qy 3841 AACATTAGAGATGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 3900
Db 4182 AAUAUACAAACUAGAGGCGAACCGUUAACACGGGGAACUUAUACUUAUCUUAUUA 4241
Qy 3901 GGGCAATTCCTCGCCGATGAGGGCTGTGCGGGCGCGCCCTACGACATCATATGTGAT 3960
Db 4242 GCGCAUUNUUGCAGUAGGCGUGUAGCCGUGGCUUAUAUUAUUAUUAUUGGAC 4301
Qy 3961 GAATGCCATGCCGTGACCTTACCACTATCTTGGCATCGGAACAGTCTTTGATCAAGCA 4020
Db 4302 GAUUGCCAUUACAUUGAGCGUACUACAUUCUUGGCAUUGGAAACAGUCCUUGACCAAGCU 4361
Qy 4021 GAGACAGTGGGTGAGACTTAACGTGTGCTGCTACAGACTAGCGCCCTCGGTGAGTGA 4080
Db 4362 GAGACCGCAGGCGUACAGGCUAGUGUUGGCGCACAGCCAGCUCUCCGGAUAGGUAACA 4421
Qy 4081 ACCCCCAACCAACATAGAGAGTGGCCCTTGGGAGGAGGCGAGATCCCCTTCTAT 4140

Db 4422 ACUCCCCACAUAAUAUAGAGAGUGGCCUUGUACAGAGGCGAGAUCCUUNUUAU 4481
Qy 4141 GGGAGGGCGAATTCCTCTGTCTTACATCAAGGAGAGAAAGATCTTCTTCGCAATCA 4200
Db 4482 GGGAAAGCUUUCCTCCAGUGUUUAUCUAAAGGGGGGAGACAUAGAUUUUUGCAUUA 4541
Qy 4201 AAAAAAAGTGTACGAGCTCGCGCGGCCCTTGGGGGTATGGGCTTGAATCAGTGGCA 4260
Db 4542 AAABAAGAGUGGACACAGCUCGACGCGCCUCCGGGGCAYGGGUGUUAUUGCGUGUA 4601
Qy 4261 TACTACAGAGGTTTGAAGCTCTCTGTATATCAACTTACAGGAGACGTATGTGTGCC 4320
Db 4602 UACUUAAGGGGUCUACAGCUCUCCGUUAUACCAACUAAAGAGACGUGUGUGUGGCC 4661
Qy 4321 ACCGACGCTCTATGACAGGATTACTGGGGAATTTGACTCCGTATCGACTGCAAGTA 4380
Db 4662 ACUGAUGCCCUAUVAGCUGGGUACCGGAGACUUGUACUCYUACUAGCAUGAAUGU 4721
Qy 4381 GCGGTCACTCAAGTTTGAAGCTTCAAGTTTGAAGCCCAATTACCAATCAACAGATT 4440
Db 4722 GCAUGUCUCAGUUAUUGUAGCUUACAGCUAGACCAACCUUACAUACCAUCAAACC 4781
Qy 4441 GTCCCTCAAGACGCTGTCTTCAAGTACAGCGCGCGGAGTGCACGCGGTAGAGGAGCTG 4500
Db 4782 GUCUCUACAGACGCUUGUCUCCGUAGUCAAAGUAGAGGAGAACUGGAGGGCGAUUG 4841
Qy 4501 GGCATTTATATGATATGTTTCACTGTGAGAGCGCTCAGAGAAATGTTTGAACATGTATG 4560
Db 4842 GGGRUUACAGAUUAUUGUAGAGGAGGCGUUGGAGUUGUACACAGUAGUG 4901
Qy 4561 CTCTGTGAGTGTACACAGCAGAGGCGCGAGTGTATGAGTCAACATCGAGACACAC 4620
Db 4902 CYUGGAGAGUAGUUAUUGCGGGGACAGCUGGUUAAGAGUUAACUUGUAGACUAGC 4681
Qy 4681 GTCAGGCTCAGGGCGATTTTCAACACGCGCGGTTTGTCTGTGCAAGACATCTTGAAG 4680
Db 4962 GUGAGACUCGCGGCGYUUAUUAACAGCCGAGUUGCCGUUAAGAACACACUGAG 5021
Qy 4681 TTTTGGAGAGGACTTTTCAACCGGCTCACAACATAGAGGCCACTTCTTTCCAAACA 4740
Db 5022 UUCUGGAAGCGGUCUUAUCAGGUCUACACACAUURACGCCCAUCUUCUCCGAGAGC 5081
Qy 4741 AAGCAATCGGGGGAATTTTGCATATCTTAACAGCTTACCAAGGCTACAGTGTGCTAGG 4800
Db 5082 AAGCAAGAGAGAGAAACUUGGCRUUAUCUAGGCGCUUACAGGCAACAGUUAUCGCCAG 5141
Qy 5142 GCAAAAGCCCUUCUUCUGUGGACGUAUGUAGUUGUUAUAGGCUCAAACU 5201
Db 4861 AACCTGTGGGCCCCACACTCTCTCTGTACCGTTGGGCTCTGTACCAACAGGTCAAC 4920
Qy 5202 ACACTUACUGGUCGCAACCCCGUUCUGUACCGUUGGUGGCGGUAACAUUGAGUVAAC 5261
Qy 4921 CTCACACATCCCGTGAAGAAATTCATCGCAACCTTGATGACGACGACCTTGAAGTATG 4980
Db 5262 UUGACGCAACCGGUGAGAAUUAUACGCGCAAGUGGACUAGCUCUAGAUUAUG 5321
Qy 4981 ACCAGACATGAGTCTTGGACAGGGGAGTCTTGGCGGCGCTGACCTTATTCGCGG 5040
Db 5322 ACAAGCUCUAVGGUCUGGCGGGGGGUGUAGCGCGCGGCAACUUAUCUGCGG 5381
Qy 5041 ACCGGTGTGTGTGATCATCGGCGGCTTGCACATTAACAGGAGACCGTCTTGGCGG 5100
Db 5382 ACUGGCUAGUUCUUAUUGGCTGCGUACACUUAUAGUUGGUGUGUGRCCCY 5441
Qy 5101 GACCAAGAGGCTCTATATGAGGCTTTTGAATGATGAGAGAAATGTCTTACGAGCGCT 5160
Db 5442 GACAAAGAAUUAUUAUAGGCUUUGUAGUAGUAGAAUUGCGCUCUCAAAGCGCCG 5501
Qy 5161 CTCATTGAAGAGGAGGAGGAGTACCGAGATGTGAAGTCCAGATTCAGAGCTTATG 5220
Db 5502 CUUAUUGAAGAGGAGCGAGUUGGCGGAGUUCUCAAUUAUUAAGGACUUCUA 5561

QY 5221 CAGCAAGCTTCAACAGCTCAAGACATACCACTGTGACGGCTTCAATGACCCAG 5280
DB 5562 CAACAGGCGCAAGAGGACGUCUACAAGCAGCAGCAGUACAGUCACUACUGGCCAG 5621
QY 5281 GTAGAACAAATCTGGGCGCAACACATGTGTGAACTTTCATAGCGGATCCCAATACCTGCA 5340
DB 5622 CUGAACAAUUUUUGGCGCAACACAUUGGAACTUCACAUUGGUUUAACAGUACUUGCA 5681
QY 5341 GGAATATCAACCTGCGAGGAAACCTGAGTAGCTTCCATGATGAGGTTCAAGTGGCCG 5400
DB 5682 GGAUCUCUCCACCTCAACCGGAAAUUCUGACAUAGUAGUAGUAGUAGUAGUAGUAG 5741
QY 5401 CTACCAAGTCCGCTGTCAACCAAGCACTATCTCTCAACATTTTGGGGGCTGCTA 5460
DB 5742 CUGACUAGCCCACTAACCCACAGCAACCAUCCUUGAACAUCAUGGAGGAGUUGUG 5801
QY 5461 GCATCCCAATTTGACCAACCGCGGGGGGCACTGGCTTCCGTTGCAATGAGCTTATGAG 5520
DB 5802 GCCUCYCAUUGGCCCCCGCGAGCCACAGCCUUGUUGUUGUUGUUGUUGUUGUUG 5861
QY 5521 GCTGCGGTAGGCAAGTATGAGCTTATGAGTATGAGTATGAGTATGAGTATGAGT 5580
DB 5862 GCGGCGGCGGAAAGAUAGGCGUUGUUAAGAUACUGUGAGCUGUUGGCGGGUAGCGC 5921
QY 5581 GCGGCGCATTTTGGGGGCTCTGCGTCATTCAGATCATGTCTGCGAGAAAGCCCTGATG 5640
DB 5922 GCAAGCAUUCAGGGGCGCUCGUAUCUUAAGAACUAGCGCGAGAAAGCCACAGUA 5981
QY 5641 GAGATGTCTCAACTTGTCTGCGGAAATTCGTCTCGGGGTGCTTGTGTGTGTGTGT 5700
DB 5982 GAAAGCGUUGGAAUUCUCUCGUCUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 6041
QY 5701 ATTCGCGGCGCATTTCTGCGCGCACTGTGGAAACCGGGGAAAGCGCGCTCAATGATG 5760
DB 6042 AUCUUGGAGCAUUCUGCGCGCGCAACGUCGUCAGGAGAGGAGCGUCUAGUGAUG 6101
QY 5761 AATAGACTATTTGCTTTTCTTCCAGAGAAATCACTGCGCCCACTACCTATGATG 5820
DB 6102 AACAGACUAGCGCCUUGCGCUCGAGGAGAAACCAUGUUGCCCUACCAUAGUGUG 6161
QY 5821 GAGTGGAGTGGTCCGACGCGTGTGACCACTATGCTGCTCCCTTACATTAACAGGCTG 5880
DB 6162 GAGUCUAGCGUCUACAGCGUGURACGAGGUGCUAGUUCACUUAUUAACAGCUUA 6221
QY 5881 CTCAAGAACTCCCAACTGATTAAGTGTAGGACTGCGCCATCCATCCATGCGGCTGCTG 5940
DB 6222 CUAAGAGACUAACUGCCUGAUCACUAAGAUUCCARUCCAUUGUCGGGUCUUGG 6281
QY 5941 CTCGCGATGTGTGGACTGGGTTTGACCACTCTTAACAGACTTTAAATTTGCTGAC 6000
DB 6282 CUCGAGACAUUUGGAGUUGGUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 6341
QY 6001 TCCAAATTAATCCCAAGATGCGCGCTCCCTTTGTCTCTGTGCAAAAGGGTATGAG 6060
DB 6342 UCAAAAUUAUCCCAAGAGUGCCCCGCAUUCUUAUUCUUGCAAAAGAAUACAAG 6401
QY 6061 GGCCTGTGGGCGGCACTGTGATGACCAACAGGCTCTTGGCGGCGCAATATCTCT 6120
DB 6402 GGUUGAUUGGCGUGUACGGGUGUUCAGUACUUCURURCCAUUGUGAGCAAAACUUCUG 6461
QY 6121 GGCATATGTCGCTTGGCTTCATGAGATCAACGGGCGCTTAAGACTGTGATATATCTGG 6180
DB 6462 GGCACUUGCCGAGUGGCGCAUAGAAAUAAACAGGCGCGAAGACUUGCUUUAACUUG 6521
QY 6181 CAGGGGACCTTTCTTCAATTTGTTACACGAGGGCGAGTGTGCGCAACCGCGCGCA 6240
DB 6522 CAGGGGACUUCUCCCAUUAUUGUUAACAAGAAAGGCGCTUGCGUGUACCCCGUCUC 6581
QY 6241 AACTTAAGTGCACCTGAGAGGTGGCGGCTCAGAGTACGCGAGGTGACGACGAC 6300
DB 6582 AAUUAACAAGACCGCAUUGGAGGUGGCGCGUCGAGUACUUGUAGUUCACACAGAU 6641
QY 6301 GGGTCAATCACTATTAACAGAGACTCACCTGATTAAGTAAAGTCCCTGCCACTA 6360
DB 6642 GGUUCUUCUUAUGUUAACGGGUAACAGGUAACACUUAAGGUCUUGCCAGGUA 6701
QY 6361 CCGTCCCGGAGTCTTTTCTGGGTGACGAGAGTGCAGATTCATATGTTTCCCCACA 6420
DB 6702 CAGCUCGAAAUUUUUCUUGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6761
QY 6421 CCGAAAGCTTTTCCGGAGATGAGTCTGTTCTGCGTTGGGCTTAATTCATTTTCTGT 6480
DB 6762 CAGGUCUUCUUCUGGAGUAGGUAAGUACGUGAGCCUUAACUUCUUGGUGUC 6821
QY 6481 GGGTCCCACTTCTTGGGACCTTGAACCCGACACAGAGTATTTGATGCTTCAATA 6540
DB 6822 GGUUCUACUUCUGGAGUUCUGAGCGGACACCGAGUACUGGCGUCUUAUGUAGCA 6881
QY 6541 GATCATCTCATATCAACGCGAGATGTGAGCGGCGGTTTATGCGGCGGAGTCAACCCCA 6600
DB 6882 GACCCGUCCAUACACGCGKAGGCGGACGACGAGCGAUUGGCAAGGAGAUUCUCCCT 6941
QY 6601 TCCGAGCAAGCTCTCGGCGAGCCAGCTATCGGACCATCGCTGCGAGCCACTGACCC 6660
DB 6942 UCAACAGUAGGUCUACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7001
QY 6661 ACCCAGCGAAGCTTATATGATGTGACATGTGTGATCTTAACCTGTTCAATGGGGCGAT 6720
DB 7002 ACCCAUAGAAGCAUUAUUGUACAUUGUAGUAGUAGUAGUAGUAGUAGUAGUAG 7061
QY 6721 GTGACTCGGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 6780
DB 7062 GUGATCCGAGUUGAGUCUACUUAAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG 7121
QY 6781 GAAGAAAGAGCGACTTTGAGCTTGTGATACATCAAGATACATGCTCCCAAGAGAG 6840
DB 7122 GAGUAGAGAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG 7181
QY 6841 TTCCCAACAGCTTACCAGCTGTGGGAGCGGCTGATTAACCAACCGCTTGTGAAATG 6900
DB 7182 TTCCCAACAGCTTACCAGCTGTGGGAGCGGCTGATTAACCAACCGCTTGTGAAATG 7241
QY 6901 TGGAAAGAGCGAGTAAACCAACCGGCACTGTGGGCTGTGATCTCTCTCTAG 6960
DB 7242 UGAGAGAGCGGAGCUAUAACCAACCUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG 7301
QY 6961 AAAACCCGACGCGCTTCCCAAGAGAGCGCGGACAGTGGGCTTAAGTGAATTCATTA 7020
DB 7302 CAAGCGAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG 7361
QY 7021 GAGATGCTTCAACAGCTGTGATTAAGTCTTTTGGCAGGCGCCCGCAAGCGGCT 7080
DB 7362 GAGGAGGAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG 7421
QY 7081 TCAGGCGCTTCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7140
DB 7422 UCCGAGUACUACUAGAGGAGGAGUACCGAGAGAGCAUUGUCCAGAACCUUCUAGCGAG 7481
QY 7141 TTGGCCCTTTCGAGAGCAAGTTCATCTCTTCATGCCCCCTCGAGAGGAGGAGT 7200
DB 7482 ACUGCGGCUCAAGAGCGGAGUACUUGUCCUAGUAGGAGGAGGAGGAGGAGGAGG 7541
QY 7201 GATTCAGACTGTGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7260
DB 7542 GACCCYGAACUUGAGUUAACAGAGGAGUAGUAGUAGUAGUAGUAGUAGUAGUAG 7601
QY 7261 GCTCCGCGCTCGAGCTCGGAGTCTGTGATCTGTGCTGAGAGGAGGAGGAGGAGG 7320
DB 7602 GUCAUUGAUUGGAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG 7661
QY 7321 TGTGCTCATATCTTCAATCTGAGCGGAGCTTAAATCTCTTGTATGCTCCGAGAG 7380
DB 7662 UGUCGUCUUAUUGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG 7721
QY 7381 GAGAAATTAACGATTAACCCCTTGAAGCACTCCCTGTGAGATATCAACAAGAGTGTAC 7440

```

Db      7722 GAGAAAGUACCGGUAUACCCUUGAGUAAUUCGUCAUGCGGUCCAAUAAAGUUGAC 7781
Qy      7441 TGTACCAACAAGAGCGCTCACTAAGGCTTAAAAAGTAACTTTGATGATGCA 7500
Db      7782 UCCCAACCCCGAGAGUUGCUUCUAGGCAAAAGAGUAGCUUUGACAGGUGCAG 7841
Qy      7501 GTGTGCACTCTCTACTGACTGACTGCTTAAAGGCAATTAAAGGAGGCTCCAGGTC 7560
Db      7842 GUGCUGAGCGCACUAGUACUAGUCUUGAGAGAGCUUAGCGGCGCCUCAAAGSU 7901
Qy      7561 ACCGAAAGGCTCTCACATGAGAGAGGCTTGCAGTTAAACCCCACTTTGCAAG 7620
Db      7902 RGUCGAGGCUCCUACAGUAGAGAAAGCCUGCGCGUAGCCCGCCCAUCCGCCAAA 7961
Qy      7621 TCTAAATATGGGTTTGGGGCTTAAGAGGTCGCGACTTTGCGGAGGCGCTTAAC 7680
Db      7962 UCCGGAUACGGAUUGGGGCAAAAGAGUUGCGCAGCUUUAUCAGAGAGGCGUUA 8021
Qy      7681 ATCAAGTCCGTGTGAAGAGACCTCTGAGAGACTCAGAAACACCAATTTCCCAAC 7740
Db      8022 AUCGCGUCCGUGUGGAGAGACUUCUGAGAGACCAACRUACCCCAUUGACACUAUC 8081
Qy      7741 ATGCGCAAAATAGAGGTGTTCTGCGTGAACCCCAACAGGGGGCAAGAAAGAGCT 7800
Db      8082 AUGGCUAAAUAUGAGUGUUGUUGCAUUGAUCUAAAGUGGGAUAAAGCCAGCUGCG 8141
Qy      7801 CTTATCGTTTACCTGACCTGCGGCTGAGGCTCTGCCAAGAGGCGCTTTATGACAT 7860
Db      8142 CUCAUCCUUAUACCCCGACCUUGGGGUCAAGGUGUGCAAAAGAGGCGCUUUAUGA 8201
Qy      7861 ACACAAAACCTTCTCAGGCGGTGATGAGGGGCTTCTATGGAATTCAGTATCCCG 7920
Db      8202 KCACAAAAGCUUCCCAAGCAUUAUAGGCGCACUUCUAGGUCUUAUACUCCCGCA 8261
Qy      7921 CAGCGGTAGAGTTTCTTGAAGAGATGAGCGGAAAGAAAGAGACCTATGGGTTTTC 7980
Db      8262 GAACGGGUCAGUUCUUCUCAAAGCUUGGGAAAGUAAAGAGACCCAAUGGGUUCUG 8321
Qy      7981 TATGATACCCGAGTCTTGAAGTCAACCGTCACTGAGAGAGACATCAGAGCTGAG 8040
Db      8322 UAUGACACCCGCGCUUUGACUCAAACGUCACGAGAGGAGCAUAAACAGAAAUCC 8381
Qy      8041 ATATATGAGGCGCTGCTCTTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8100
Db      8382 AUAUAUAGGCUUUGUUCUGCCUCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 8441
Qy      8101 AGACTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 8160
Db      8442 AGACUUAUAGUAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 8501
Qy      8161 TCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8220
Db      8502 UGCAGCGCAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8561
Qy      8221 GCCTTAGCGGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8280
Db      8562 GCCCUUAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 8621
Qy      8281 GACTTGTGTCTCTCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8340
Db      8622 GACUUGUUCUUAUAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8681
Qy      8341 TTCAAGGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8400
Db      8682 UUCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 8741
Qy      8401 GATGTGAGGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8460
Db      8742 GACUUGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8801
Qy      8461 CGCGGAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8520

```

```

Db      8802 CGCGCGGAGUACUUCUUAACAGAGACCCUACACUCCAAUACCCGAGGUCUGGAA 8861
Qy      8521 ACAGTTAGACATCCCTGTCAATTCAATGCTGAGGAAACATATCAAGTACGCCCCGAC 8580
Db      8862 ACAGUAAAGACUCCCGUACAUAUUCUGGCGGAGCAACUACUAGUACGCCCCACA 8921
Qy      8581 ATATGAGGCTGCAATGCTGATGAGACACATTTCTTCTCAATTTCTATGAGTCAAG 8640
Db      8922 AUCUGGUCGAGUAGGAGUACAUAUAGCUCUUCUACUACUUAUUGGCGAGACU 8981
Qy      8641 CTGACCAAGAACTCAACTTGAAGATGACGAGCGGTATCTCTGAGTCCCTTGAC 8700
Db      8982 CUGAACCAAAUUCUUAUUGAGUAGUAGGCGAGUUAUCUGGUCUUAUUGAG 9041
Qy      8701 CTCCAGCTTAAATGAAAGTTACATGAGGCTTGAAGCTTTTCTGCAACATAC 8760
Db      9042 CUACGCGCAUUAUUGAAAGCUAUGGAGGAGUAGGAGGAGGAGGAGGAGGAGGAG 9101
Qy      8761 CCCACGAACTGACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8820
Db      9102 CCCACGAAUUCUACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9161
Qy      8821 GCGTGAAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8880
Db      9162 GCGUGAAAGAGUUGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9221
Qy      8881 GCGGTTGGGAGTATCTCTTCAATGAGGCGGTGAAGCAAGCTCAACTCA 8940
Db      9222 GCAUUGUGGCGGCUACUUCUUAACUGGCGGAGUAAACAAAGCUCAUACUCCA 9281
Qy      8941 TTGCGGAGAGCGGCTCTCTGATTTATCAAGCTGATTTCAAGCTGAGGAGGAGGAG 9000
Db      9282 UUGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9341
Qy      9001 GACATTTATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9060
Db      9342 GACAUUAUACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9401
Qy      9061 CTTTGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9099
Db      9402 CUUAGCGUAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9440

RESULT 5
US-09-827-688-6
; Sequence 6, Application US/09827688
; Patent No. 6821955
; GENERAL INFORMATION:
; APPLICANT: ORSON, FRANK
; APPLICANT: KINSEY, BERNARD
; APPLICANT: BHOAGAL, BALBIR
; TITLE OF INVENTION: MACROPHAGE-INDUCED PROTEIN CONJUGATES AS ORAL GENETIC IMMUNIZATION
; FILE REFERENCE: P01949US1/10004014
; CURRENT FILING DATE: 2001-04-06
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 9413
; TYPE: DNA
; ORGANISM: HEPATITIS C
US-09-827-688-6

Query Match 45.5%; Score 4141.8; DB 4; Length 9413;
Best Local Similarity 66.6%; Pred. No. 0;
Matches 606; Conservative 0; Mismatches 2952; Indels 93; Gaps 6;

Qy      1 ATGAGCAAAATCTTAAAGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60
Db      330 ATGAGCAAAATCTTAAAGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 389

```

QY 61 GACGTTAAGTTCCGCGCGCGCGCGAGATCGTTGGCGAGATATCTTTGGCCGCGCAGG 120
Db 390 GACGTTAAGTTCCGCGCGCGCGCGAGATCGTTGGCGAGATATCTTTGGCCGCGCAGG 449
QY 121 GGGCCGAGGTTGGGTGGCGCGCGCAAGAAAGATTGGAGCGGTCCAGCCAGTGA 180
Db 450 GGGCCGAGGTTGGGTGGCGCGCGCAAGAAAGATTGGAGCGGTCCAGCCAGTGA 509
QY 181 AGGCGCGAGCCCATCCCTAAAGATCGGCGCTCACTGGCAAACTCTGGGAAAAACAGA 240
Db 510 AGGCGCAACCTTATCCCAAGGCTCGCGCCAGAGGTAGAACTTGGACTCAGCCGCGG 569
QY 241 TACCCCTGGCCCTATACGGAATGAGGGAATCGGCTGGGCAAGATGGCTCTGCCCC 300
Db 570 TACCCCTGGCCCTATATGCAACGAGGGTATGGGGTGGCAGAGATGGCTCTGTCACC 629
QY 301 CGAGGTTCCGCTCCTCTTGGGGGCCCAATGACCCCGGCAATAGGTGGCGCAAGTGGGT 360
Db 630 CGTGGCTCTGGCCCTAGTTGGGGGCCCAAGACCCCGGCGTAGGTGGCGTAATTTGGGT 689
QY 361 AAGGTCATGATACCCCTAAGCGGCGCTTGGCGAGCTCATGGGGTACATCCCTGTCGTG 420
Db 690 AAGGTCATGATACCCCTTACATGCGGCTTGGCGAGCTCATGGGGTACATCCCTGTCGTG 749
QY 421 GGGCGCCGCTCGGCGCGCGCGCGAGAGCTCTGGCGCATGGCGTGAAGTCTTGGAGAGAC 480
Db 750 GGGCGCCGCTAGGGGGCGGTGCGCAGGGCCCTGGCGCATGGGTGTCGGGGTTCTGGAGAGAC 809
QY 481 GGGGTTAATTTTGGCAACGAGGAATTACCGGGTGTCTCTTTTCTATCTTCTTGGTGGCC 540
Db 810 GGGGTTAATTTTGGCAACGAGGAATTCGCGGGTGTCTCTCTATCTTCTTGGTGGCC 869
QY 541 CTGCTGTCTGTGATACCAACCGGGTCTCGGCTGGCGGAAGTGAAGAAATCATGATACGGGC 600
Db 870 TTGCTGTCTTGTGATACCAACCGGGTCTCGGCTGGCGGAAGTGAAGAAATCATGATACGGGC 929
QY 601 TACATGTGATTAACGACTGCAACCAATGACAGATTTACCTGACAGCTTCAGGCTGCTGTC 660
Db 930 TACATGTGATTAACGACTGCTGCCAATCAAGTATGTATGAGGCGAGCGGACATGATC 989
QY 661 CTCCAGTCTCCCGGGGTGGTCCCGTGGCGAAGAAATGGAGAAATGATCTCAAGTCTGAGTA 720
Db 990 ATGCAACACCCCGGGGTGGTCCCGTGGCGAAGAAATTTCTCCGCTGGTGGAGTA 1049
QY 721 CGGCTCTACCGAATGTGCGCGTGGCAGCGCGCGCGCGCTCACAGAGGGCTTGGAGAG 780
Db 1050 GGGCTCACTCCAGCGCTGGCGCGCGAGAAACAGAGAGATCCCAACAGCAATACGAGGC 1109
QY 781 CACATGACATGTGTGTATGTCCGCAAGCTGTGCTGTGCTCTTACGTTGGGGAGCTC 840
Db 1110 CACGTCGATTTGTGCTGTGGGGGGGCTGCTCTGTGTCCGCTATGTAGTTGGGGATCTC 1169
QY 841 TGGCGTGGGGGTGATGTCTGAGAGCCCAATTTTCAATTTGTCTGGCGGACACCATGTTTT 900
Db 1170 TGGCGATCCGTTTCTGTCTGTCCAGCTGTTCACCTTCTCACTCGCGGATTAAGAGCG 1229
QY 901 GTCAAGAATGCAATTTGCTCCATCTACCTGGTACATCACTGAGACACCGATGGATGG 960
Db 1230 GTCAAGAATGCAATTTGCTCAATCTATCCCGGCGCGATCAAGGTACACCGATGGCTTGG 1289
QY 961 GACATGATGATGAATGTGTGCGCCACGGCTACCATGATCTTGGCTAGCGCATGTGTCTC 1020
Db 1290 GATATGATGATGAATGTGTGCGCCACGGCTTACATGATGATGATGATGATGATGATGATG 1349
QY 1021 CCGGAGGTCATTAATGACATCTTAAGCGGGCTCATTTGGGGCGTATGTTGGCTTGGCC 1080
Db 1350 CCGAAGGCGTGTGACATGATGAGCGGGGCCCATGGGGTGTCTTAAGCGGGCTTGGCC 1409
QY 1081 TACTTCTCTATGACAGGAGCGTGGCGAAGTGTGTCTCTTCTGTGGCGCGCGGG 1140
Db 1410 TACTTCTCTATGATGGGAGACTGGGCTAAGGCTTGTGATGATGATGATGATGATGATGATG 1469

QY 1141 GTGAGCGCGGACCCCATCTATTGGGGGTGTTCGCGCGAGACCAACCGGCGCCTCACCC 1200
Db 1470 GTTGAACGGGACACCCCATCTGACAGGGGGAAGGGTGTAGCTTCCAGACACCCAGAGCTCGTG 1529
QY 1201 AGCTTATTTGACATGGGCCCCGAGGAAATTCAGCTGTGTTAAACCAATGGCAGCTGG 1260
Db 1530 TCGTGGCTCTCAACAAAGGCCCATCTCAAGAAATTCAACTGTGAACACCAAGCGAGCTGG 1589
QY 1261 CACATCAACCGGACCGGCTGAACTGCAATGACTCTTTGACACCGGCTTTATTCGGTCT 1320
Db 1590 CACATCAACCGGACCGGCTGAAATGCAATGACTCTTTGACAACTGGTTATTTGGTGGG 1649
QY 1321 CTGTTCTAACCCACAGCTTCAACTGTGAGATGTCCTGAAACGATGTCCGCTGGCC 1380
Db 1650 CTGTTCTAACCCACAGCTTCAAACGCTGTGGGTGTCAGAGCGCATGTGCTAGCTGGCC 1709
QY 1381 AGTATGAGAGCCCTTCCGGGTGGGATGGGGGCGCTTGCAATATGAGGATTAATGACCAAT 1440
Db 1710 CCGATGATGATGATGCTGCTCAGAGGGTGGGTCCCATCACTCATGATATGCTTGAAGAGCTC- 1768
QY 1441 CCAGAGATATGAGACCTTATTGCTGGCACTACCCACCAAGGCGAGTGGCGGTCTCC 1500
Db 1769 -----GACACCAAGAGCATATTTGCTGGCACTAGCGGCTTCGACCGTGGCGGATCGTGCT 1823
QY 1501 GCGAAGACTGTGTGGGCCCAAGTATCTGTTTCAACCCCAAGCCAGTGTAGTGGGACG 1560
Db 1824 GCGTCGCAAGTGTGTGGTCAAGTATGTTGCTTCACTCCAGCGCTGTTGTAGTGGGAGG 1883
QY 1561 ACCGACAGCTTGGAGCGGCGCACTTACAGTGGGGGGAATGAGACAGATGCTTCTTA 1620
Db 1884 ACCGATCGTTTGGGCGCTCTCAAGTATAGCTGGGGGAATGAGACAGATGCTTCTTA 1943
QY 1621 TTGACAGACTTGAACCAACCGCTGGGGTATGTTTGGCTGGCAGTGAATGAACTTTCT 1680
Db 1944 CTTAGCAACACGGGCGGCGCTCAAGGCACTGTTTGGGTGACGATGATGAACAGCACT 2003
QY 1681 GGTACACCAAGACTTGGCGCGCAACACCTGCGCTGATTAAGAGCTGATTCAAACGCGAC 1740
Db 2004 GGGTTTACCAAGACTGTCGGGGGCGCTTCCGTGACATCGGGGGGTCCGCAACA- 2061
QY 1741 ACGGACCTGTTGGGCCCAAGCGCTGTTTGAAGAACCTGATACCACTTACCTCAAA 1800
Db 2062 ----CTTGTGTGCCCCACAGATGTCTTCGGAACACCCGAGGCCCTTACACAAAG 2117
QY 1801 TGGGCTCTGGGCGCTGGCTGCAAGCAAGTGTCTGATGACTACCTTACAGGCTCTGG 1860
Db 2118 TGTGGCTGGGCGGCGCTGTTGAACCAAGGTGATGTTGATACCATACAGGCTCTGG 2177
QY 1861 CATTACCCCTGCAAGTTAATATACATCTTCAAAATGAAGATGTATGTGGAGGGATT 1920
Db 2178 CACTTACCCCTGCACTGTAACTTTACCGTCTTTAAGTACAGATGTATGTGGGGGCGTG 2237
QY 1921 GAGCAGAGGCTCAAGGCTGATGCAATTTCACTCGTGGGGAATGTTGAACCTTGGAGAG 1980
Db 2238 GAGCAGAGGCTCAAGGCTGATGCAATTTGAATCTGAGAGAGGCTGTGAATCTTGGAGAG 2297
QY 1981 AGAGACAGAAATGCACTGTCTCTTGTGTGACTCCACCAAGCAAGATGGCAATTTTACCT 2040
Db 2298 AGGATATGATGCAAACTCAGCGCGGCTGTGTGTCTTCAACAAGAGTGGAGATACGCGC 2257
QY 2041 TGTCTTATCTGGAAGCTGCGCGCTTGTGTGATGCTGTCTTCTCACTCCACCAAAACATC 2100
Db 2358 TGTTCCTTCAACACCTTACCGGCGCTGTGCACTGGCTTGTATCTTCACTCCGAAACATC 2417
QY 2101 GTGAGAGTCAATTCATGTATGAGCCATACACCTGGCCCTCAAAATATACCTGCGATGG 2160
Db 2418 GTGAGAGTCAATTCATGTATGAGGTGAGGCTGAGAGTGTCTCTTTGCAATCAAAATGG 2477
QY 2161 GAGTGGGATATCTCTTAATCTCTGCTCTTACGCGAGCGCAAGGATTTGGCGCTGATAG 2220
Db 2478 GAGTATATCTGTGTCTTCTCTTCTTGGCGGAGCGGCGGTCTGTGTGCTGTG 2537
QY 2221 ATGCTCATCTTGTGGCGCAGGCGCAAGACGACTAAGAGAGCTGTCTATCTTGGCAGCT 2280

2538 ATGATGCTGTGATAGCCAGGCTGAGGCGACCTTAGAAGACCTGGTGCCTCAATAGCG 2597
2281 GCCAGGCGAGCTAGCTGCAATGGCTTCTATATTTTGTATCTTTTTCGTGGCTCTGG 2340
2598 GCGTCTGTGGCCGAGCGCATGGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2657
2341 TACATCAAGGGTGGGAGTACCCCTTAGTACCTATTCCTCACTGGCTGTGGTCTT 2400
2658 TACATCAAGGCGAGGCTGTGCTCTGGGGCGGCATATCTCTATGGCGTATGGCGTTG 2717
2401 AGCTATCTGCTCTAGCATTTGCCCAACAGGCTTATGCTTATGACGATCTGTGCATGC 2460
2718 CTCTGCTCTGTGCTGAGCTTACCAACGAGCTTATGCCATGAGCCGAGATGGCTGCA 2777
2461 CAAATAGAGCGGCTCTGCTGTGTATGATCATCTCTTATCTCTACCCCCGGGTATAG 2520
2778 TCGTGGAGGCGCGGTTTTTGTAGGTCTGTACTCTTGAACCTTGTACCATATCATAG 2837
2521 ACCCTTCTGACCGGTTTTTGTGGTGTGTGCTATCTTGAACCTTGGGGGAAGCTATG 2580
2838 GTTCTCTGCTAGGCTCATATGTGTGTTCAATATTTATCACAGAGCCGAGGCCAC 2897
2581 GTCCAGAGTGGGCGACCACTATGCAAGTGCAGGTGGCCGTGATGGCATATATGGGCC 2640
2898 TTGCAAGTGTGGTCCCGCTCTCAATGTTGGGGAGGCGCGATGCAATCTCTCTT 2957
2641 GTGCGCATATTTCAACCAAGTGTGTGTGATCAATAACAAATGCTCTTGGCGGTCTT 2700
2958 ACATGGCGGCTCATCCAGAGCTATCTTTGACATCAACAACTCTGCTGCGCATATCTC 3017
2701 GGGCTGCTTACCTCCATAAAGGCTTTGACCGGCGTGGCGTACTTGTGACAGGCTCAC 2760
3018 GGTCCGCTATGTGTCTCCAGGCTGCGATATCTAGAGTGCCTGATTTTGTACGCGCTCAG 3077
2761 GCTCTACTGAGATGTGCACATAGGCAAGGATCTCGCGGGGGCAGGTACGTCCAGATG 2820
3078 GGGCTATCGGTCATGATGATGTTAGTGCAGAGGTGCTGGAGGCCATATGTCCAAATG 3137
2821 GCGCTACTAGCCCTTGGCAGGTGATCTGGCATTTATCTATATCAACCTCAACCCCTATG 2880
3138 GCTTATGATGAAGCTGGCGCGCTGACAGGTACTATATATGACATCTTACTCCACTG 3197
2881 TGGATTGGGCTGTGATGGGCTGCGGGAGCTGGCGGTGGCGTGAAGCTTATCATCTTC 2940
3198 CGGATTGGGCCCAACGCGGCTCTACAGACCTTGGCGAGTGAAGCCGCTGCTTC 3257
2941 AGTCGATGAGAGAAAGTCAATTGTCTGGGAGCGGAGACAGCTGCTTGGGAGCATT 3000
3258 TCTGATATGAGATTAATCTCATCACTGGGGGAGAGACACGCGCGGTGTGGGAGATC 3317
3001 TTACACGAGTCTCCGTGTCCGCCCACTTGTGTGGAGGTCTCTTGGCCCACTGAT 3060
3318 ATCTCGGGTCTACCAATCTCCGCCCAAGGGGAGAGATTACTTCAAGAACCGGCGAT 3377
3061 GGTATATCTCCAAAGGGGTGAGATCTTTCGCCCCCATCACTGCTTACCGCCAGAGACA 3120
3378 AGTTTGGAGAGCAAGGGGTGAGGCTCTTGGCTTATCAAGCTATTTCCAAACAAAG 3437
3121 CGTGGCTTTTGGGCAACATAGTGTGAGCATGACGGGCGCGCAAGACAGAGAGCT 3180
3438 CGGGGCTGCTTGGCTGTATCATCACTAGCTCAAGGTGCGGACAAAGACAGGTGAT 3497
3181 GGGGAATTCAGGTCTGTCTCAAGTCACTCAAGTCTTCTCTGAAACATCATCTCTGGG 3240
3498 GGGGAGATTCAAGTGTCTCTCAACGCAACCAATCTTCTGCGAGCTGCTCAATGTGC 3557
3241 GTTTTGTGACTGTCTACATGAGCTGAGCAACAAGCTGTGGCGGCTCACGGGGTCCG 3300
3558 GTGTGTGACCGTCTTACATGTGTGCGGCTCGAAGACCTTGGCGGCGCCGAAGGATCA 3617
3301 GTACGCAATGTATCTCAAGTGTGAGGGGAACTTATAGGTGTGCCCAAGCCCTCTGG 3360

3618 ATCACCCAATGTATCAACATGTAGACCAAGACCTGTGGTGGCGGCGCCCGGG 3677
3361 ACTAAATCTTTGAGAGCGGTGCAAGTGTGAGCGGTGACTGTGACTTGTGCAAGCGGAAC 3420
3678 GCGGCTTCATGACACCTGTGCACCTGCGAGCTCGAGCTTTACTTGTGTCAAGAGCAT 3737
3421 GCTGATGTCAATCCCGGCTGAGAGACCGGGGACAAAGGGGAGCGCTACTCTCCGAGA 3480
3738 GCTGATGTGTTCCGGTGGCGCGGGGGAGACAGAGGGGAGCGTCTTCCCGAAG 3797
3481 CTTCTTTCACTTGAAGGGGTCTCAAGAGCCCGGTGTATGCCCCAGGCGCACGCT 3540
3798 CCCATCTCTTACCTGAAGGGGCTCTCGGGTGAACCACTGCTTGGCCCTTGGGGCGT 3857
3541 GTGAGGTCTTCCGGGAGCTGTGTGCTCCGGGGGTGGCTAATGCAATGATTTCATC 3600
3858 GTAGGATCTTCCGGGCTGTGTGACCCCGGGGGTGGCGAAGCGGTGACTTATA 3917
3601 CCGTTGAGACACTCGACATCGTCAAGCGGTCCCGCACTTTTATGTAACAACAGACCA 3660
3918 CCGTTGAGTCTATGGAACATACATGCGGTCTCCGCTCTTCAACAGCACTCATCCCT 3977
3661 CTTGTGTGCCCGACACTATCAAGTGGGTACTTGCATGCCGACGTGGCACTGGAAG 3720
3978 CCGGCGTACCGCAACATTTCCAAAGTGACATTTACAGCTCCACTGGACGGCAAG 4037
3721 AGACCAAGTTCCTGTCCATATGTGCTCAAGGGTATTAAGTGTATGTCTTATCCC 3780
4038 AGACCAAGTGTCCGCTCATATGTGACGCCAAGGGTATCAAGTGTCTCTTAAACCG 4097
3781 TCAGTGTGCCACCTGTGGGTTTGGGGGTACTTGTCTTAAGGACATGACATCAATCCC 3840
4098 TCCGTTCCGCGCACATTTGGGCTTTGAGGTATATTCAGAGGCAATGGCATCGAGCT 4157
3841 AACATTAGACCTGAGTCAAGACTGTGACGACCGGGCGCCCATCACTCATCTCAATAT 3900
4158 AACATCAGAACTGGGGTAAAGACATCACACGGGGCGCCCATCACTCACTCATAT 4217
3901 GCGAAATTCCTGCGCATAGGGGCTGTGGCGGGCGCTTACGACATCATATATGAT 3960
4218 TGCATGTTCTTGCACGAGTGTGATCTCGGGGGCGCTTATGACATCAATATGAT 4277
3961 GAATGCAATGCGGTGACTCTACCACTCTTTGATCGAAGCACTCTTGTATCAAGCA 4020
4278 GAATGCACTCAACTGACTGACTGACTACCTTGTGGGATCGGACATCTCTGATCAGGCA 4337
4021 GAGACAGTGGGCTGACTTAATCTGTGTGCTGACTACGCTACCGCCCTGGGTCAGTGACA 4080
4338 GAGACGGCTGAGCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4397
4081 ACCCCCAACCAACATTAAGAGGTGGCCCTTGGGCAAGAGGCGAGATCCCTTCTAT 4140
4398 GTGCCACACCCCAACATCAAGAGAGGTGGCTGTCCAAACTGGAAGAGATCCCTTCTAT 4457
4141 GGGAGGGGATTTCCCTGTCTTATCATCAAGGAGAGAAAGATCTGATCTTCTCCATTCA 4200
4458 GGGAAAGCATCCCATTTAAGGCGATCAAGGGGAGAGGAGATCTCTTCTGCAATTC 4517
4201 AAGAAAAGTGTGACGAGTCTGGGGGCGCTTCCGGGATATGGGCTTGAATCTCAGTGGA 4260
4518 AAGAAAGTGTGACGAGTCTGGCGCAAGCTGACAGGCTCGGACTCAATGTGTAGGCG 4577
4261 TACTACAGAGGTTGAGAGTCTCCGTATTAACCACTCAAGGAGAGCTAGTGTGTCTGCGC 4320
4578 TATTACCGGGGTCTCATGTGTCTGCTATACGAGCTAGCGGAGACGTCTGTGTGTGGA 4637
4321 ACCGAGCGCTCATGACAGGATATCTGGGAGCTTTGACTCCGTGATCGACTGCAAGTGA 4380
4638 ACAGAGCTCTATATGACGGGTTTACCGGCGACTTTGACTCAAGTATCGACTGCAACACA 4697
4381 GCGGTCACTCAAGTTGTATGACTTCAAGTTTGAACCCCAATTCACATTAACACACAGATT 4440
4698 TGTGTACCCAGACAGTCAATTTCAAGCTTGTATCCCACTTGCACATGTAGACAGACAG 4757

4441 GTCCCTGAAGACGCTGTCTCACTAGCCAGCCGCGGGTGCACGGGTAGGGGAAGACTG 4500
4758 CTGCCCAAGACGGGGTGTGCGGTGCGGAGCGGAGGTAGGACTGGCAGGGGAGGAGT 4817
4501 GGCATTTATAGATATGTTTCCACTGTGTAGCGAGCCCTCAGGAATGTTTGAACATGTATG 4560
4818 GGCATCTACAGATTTGTGACTCCAGGAAAGCGCCCTCAGGACATGTTTGAATCTCTCGT 4877
4561 CTCTGTGTGTCTGACAGCGAGGGGCGCATGTATGAGCTCACACCATCGAGAGACACC 4620
4878 CTGTGTAGTCTATGACGAGCGCTGCGCTTGTATAGCTCAGCCCGCTGAGACCTCG 4937
4621 GTCAAGCTCAGAGCGGTATTTTCAACAGCCCGGTTTGTGTGTGCAAGAACATCTTGTAG 4680
4938 GTTAGGTGTGGGGCTTACCTAATACACAGAGTTGCGCGTCTGCAAGAACCACTTAG 4997
4681 TTTTGGAGGACGATTTTCAACCGGCTCTCAGACATATAGTCCCATCTTCTCCAAACA 4740
4998 TTTGTGGAGAGCGTCTTCAACAGGCTCACCACATAGTCCCATCTTGTGTCCAGACC 5057
4741 AAGCAATCGGGGAAAAATTTGCACTACTTAACAGCTACAGGCTACAGGTAGTGGCTAG 4800
5058 AAACAGCAGAGACACCTTCCCTACTGTGTAGCATACAGGCAAGTGTGCGCAAG 5117
4801 GCCAAAGCCCCCCCCGCTCTGTGGGACGTCACTGTGAAAGTGTGTGACTCGACTCAAGCCC 4860
5118 GCTCAGGCTCACACCTCATCGTGGGACCAATGTGAAAGTGTCTCATACGGCTAAACCC 5177
4861 ACACTCGTGGGCCCCACACTCTCTGTACTCGCTTGGGCTTGTTTACCAAGAGGTAC 4920
5178 ACACTGACATGGGCAACCGCCCTGTGTACAGGCTAGAGCGCTTCAAAATGAGGTACT 5237
4921 CTCACACATCCCGTAGAGAAAATACATGCGACCTGCATGCAAGCGACCTTGTAGGTATG 4980
5238 CTCACACACCCCATACCAAAATACATCATGCGATGATGTGCGCTGACCTTGTAGAGTGT 5297
4981 ACCAGCAGATGGGCTTTGGCAGAGGGAGTCTTGGCGGCGTGTGCGCGGTATTTGCTGAG 5040
5298 ACTAGCAGCTGGGTCTAGTAGAGGAGGAGTCTTGGGCGCTGTGCGGCTGACTGCTGAG 5357
5041 ACCGGGTGTGTTGATCATGCGCGCTTGCACATTAACAGCGAGCGCTGTGCGCG 5100
5358 ACAGGACAGTGTCTATGTGGGCGAGATCATCTTGTCCGGGAGCGCCAGCTGTATTTCCC 5417
5101 GACAAGAGGTCTCTATAGGCTTTTATGATGAGATGAGGATGCTTCAAGGCGGCT 5160
5418 GACAGGAGATCTCTTACAGAGATTGATGATGAGAGATGTGCTTCAACCTCCT 5477
5161 CTCAATTGAAGAGGGCAGCGATAGCCGAGATGCTGAAGTCCAGATCCAGATCCAGCTTATG 5220
5478 TACATCGAGCAAGGAATGAGTGTGCGGCAATTCAACAGAAAGGCGCTCGGATTGCTG 5537
5221 CAGCAAGCTTCCAAACAGCTCAGACATACACCCACTGTGACAGGCTTCAATGCGCAAG 5280
5538 CAACAGCGCACCAAGCAAGCGAGGCTGTCTCCGCTGTGTGATGCCAATGTGCGAGCC 5597
5281 GTAGAACAATTCTGGGCAAAACACATGTGAACTTATTAGGCGATCCAAATACCTGCA 5340
5588 CTGAGGTCTTCTGGGCGAAGACATGTGAACTTATCAGGCGGATACATGATCTTGCA 5657
5341 GGAATATCAACACTGCGAGGAAACCTGCAAGTGTTCATGATGCGCTTCAAGTGTGCGCC 5400
5658 GGCCTATCACTCTGCTGAAACCCGGAATAGCATATGATGCTTTACAGCCCT 5717
5401 CTCACCAAGTCCGCTGTCAACAGCACTATCTTCTCAACATTTTGGGGGCTGTGCTA 5460
5718 ATCACACAGCCGCTACACACCCAAATATCCCTCTGTTTAACTCTTGGGGGATGTGGTG 5777
5461 GCATCCCAAAATTGACACACCGCGGGGCACTGCTTGTGTGAGTGTGCTGTAGTGGGA 5520
5778 GCTGCCCAAGCTGCTCCCGCCAGCGCTGTGCTTGTGTGAGGCGCGGCAATTTGCCGCT 5837

5521 GCTGCCGTAAGGAGATATAGCTTATAGGTAAAGTGTAGTGAACATCTGGCAGGATATGCT 5580
5838 GCGGCGGTGTGACAGATATAGTCTCGGGAAGGATCTTGTGACATCTTGGCGGCTATGGG 5897
5581 GCGGCGATTTGGGGGCTCTGCTCCCATTCAGAAATCATGTCTGTGGGAGAGGCCCTCCATG 5640
5898 GCGGGGTGTGGGCGACATGTGTGCTTTAAGTATAGCGGAGATGCTCTCACT 5957
5641 GAGGATGTCTCAACTTGTGCTGCTGGAATCTGTCTCGGGTGTGCTGTAGTGGAGTCT 5700
5958 GAGGATCTGTTAATTTACTCTCTGCACTCTTCTCTGCGGCGCTGTGTGTGGGGTCT 6017
5701 ATCTGCGGCAATTTCTGCGGACAGCTGTGGAACGCGGGGAAGCGGCTCCAAATGATG 5760
6018 GTGTGCGGACAAATCTGCTGCGGACGTGGCGCGGAGAGAGGGGCTGTGACGTGAGTG 6077
5761 AATGACATATGCTTGTGCTTCCAGAGGAATACAGTGTGCGCCCAACCCACTACGTAGG 5820
6078 AACCGCTGATAGCGTTCGCTTGTGCGGGTAAACAGTGTCCCGCAGCAGCATATGTGCC 6137
5821 GAGTCGATGCGTCCGAGCGTGTGACCCAACTACTTGTGCTCTTACATTAACCAAGCCTG 5880
6138 GAGAGGAGCGCGCGCGGCTGTACTAGATCTCTCCAGCCTTAACATCACTAGTGG 6197
5881 CTCGAAGAATCCACAACTGTGATTAAGTGAAGTCTGCCCATCCATGCGGCGCTGTGG 5940
6198 CTGAAGAGCTTCAATGATGATTAATGAGACTGTCTCCAGCCCTTGTTCGCGCTGTGG 6257
5941 CTCGCGATGTGTGGGACTGGGGTTTGCACATCTCAACAGACTTTAAATTTGGCTGACC 6000
6258 CTAAAGATGTTTGGGACTGGAATATGACAGTGTGTGAGTCTCAAGACTTGTGCTCCAG 6317
6001 TCCAAATTAATCCCAAGATGCGCGGCTCTCCCTTGTCTCTGTCAAAAGGGATACAG 6060
6318 TCCAACTCTGTGCGGCTTACCGGACCTCCCTTCTGTCAATGCAAGCGGGTACAG 6377
6061 GCGGTGTGAGCGGCACTGGCATATGACACACAGTGTCTTGTGCGGCGCCATATCTCT 6120
6378 GGAGTCTGCGGGGGATGGCATATGACAAACCACTGTCCATGTGTGACACAGATCAC 6437
6121 GCGAATGTCGCTTGGGCTCAATGAGATACAGGGGCTTAAGCCTGCATGAATATCTGG 6180
6438 GGCATGTCAAAAATAGCTCTCATAGAGATTTTGGGCAAAAACCTGACGACACATGG 6497
6181 CAGGGACCTTCTCTATCAATTGTTACACGAGGCGCAGTGTGCTCGGAAACCCGCGCA 6240
6498 CATGGAACATTTCCCATCAACGATACACAGGGGCCCTGACGCGCTCCGAGGCGG 6557
6241 AACTTAAAGTGTGCATCTGAGAGGTGGCGGCTCAGAGTACGCGAGGTGACGACAC 6300
6558 AACTTATTCAGGGCGCTGTGTGCGGGTGTGCTGAGAGATAGGTGAGGTTACGCGGTTG 6617
6301 GGTATATCACTACATTAACAGGACTCACCACTGATPACTTGAAGTCCCTGCGCAACTA 6360
6618 GGGGATTTCCACTAGTGAAGGAGCATATGACCATGAAAGGTGAATGCCATATGCAAGTT 6677
6361 CCTCTCCGAGTCTTCTTCTGTGTGAGCGGAGTGCAGATGCATATGATTTGCCCCACA 6420
6678 CAGCCCTGAAATTTTTCACGAGAGTGAATGAGTACGATGCAAGTATGCTCCAGTG 6737
6421 CCGAAGCGTTTTCGAGATGAGTGTCTGTTGTGCTTGTGGCTTAATTCATTTGTGCTG 6480
6738 TGCAAACTCTCTTACAGAGGAGGTGTGATTTCCAGATGGGGCTCAACAGTACCTGTGCT 6797
6481 GGGTCCAGCTTCTTGTGACACCTGTAACCGGACACAGAGTATGATGTCCATGCTAACA 6540
6798 GGGTCAAGCTTCCATGTAGGCGCAACCGAATGTGCAAGTGTCACTTCACTGTCTAC 6857
6541 GATCAATCTATATCAACGCGAGACTGACGCGGCTTTTGAACGCGGGGTCAACCCCA 6600
6858 GACCCCTCTATATTTACAGAGAGACGCGCAAGCGTATAGCTGTGCAAGGGGTCTCCCCC 6917
6601 TCCGAGCAAGCTCTTGTGCGAGCCAGTATGTGCGACCATCTGCTGCGAGCCACTGTGACC 6660

Db 6918 TCCCTGGCCAGACTCTTAGCTAGCCAGATTGTCGCGCCTTCTTTGAAGGCGACATGTAAT 6977
Qy 6661 ACCCAAGCGCAAGCTTAGATGTGTGACATGTGTGAATCTAACT- - - - -GTTCC 6708
Db 6978 ACCCATCATGATCTCCCGAGCGCTGACCTCATCGAGCGCAACCTCTGTGCGGACAGAG 7037
Qy 6709 ATGGGGGCGCATGTGACTCGGATAGAGCTGTGGGTCAAAGTGTGCTTGTGGACTCTCTC 6768
Db 7038 ATGGGCGGGAACATCAACCCGTGTGAGTCAGAAAAATAGGTGTAACTCTGGACTCTTTC 7097
Qy 6769 GACCCAAATGTGCAAGAAAGAGCGACCTTGAAGCTTGCATACCATGACATGATAGTCTC 6828
Db 7098 GATCCGATTCGGGCGGTGAGAGATGAGAGGAAATATCCGTCGCGGCGAGATCTTGGA 7157
Qy 6829 CCCAAGAAAGAGTTCACCAACAGCTTTCACCGGCTGGGCGACGCTGATTTACACCCACG 6888
Db 7158 AAACCCAGGAAGTTCCTCCCGACGCTTGCATATGGGCAAGCCCGGATTTACACCTTCA 7217
Qy 6889 CTTGTGGAATCGTGGAAAGGCGCAGATTACCAACGGGCACTGTTGGGCGCTGTGCTCTC 6948
Db 7218 CTGCTAAGATCTGTGAAGAGACCCGAACTAGTCCCGCGTGTACACGGGTGCTTGTG 7277
Qy 6949 CTTCTCTTAGAAAAACCCGACGCTTCCCGCAAGAGAGCGCGCAGACAGTGGGCTTAAGT 7008
Db 7278 CCAATCTACAAAGGCGCCCGCAATACCACTTCCAGAGAAAGAGACGTTGTCTGACA 7337
Qy 7009 GAGACCTCAATAGAGATGCTCTTCAACAGCTGGCCATTAATGCTCTTGTGGCAGCCCCC 7068
Db 7338 GAGTCCACCTGTCTTGTGCTTGGCGAGCTGCTACTAAGACTTTGGGAGCTCCGGG 7397
Qy 7069 CCAAGCGGCGATTGAGGCTTTTCCAGGGGCGGGCGCTGCGATTCGGGAGTCAAGAG 7128
Db 7398 TCTGCGGCGCTTGTACAGGCG- - - - -CACGCGACTGCGCTTCCG- - - - -TGAAGCC 7445
Qy 7129 CTTCTGATGATGTTGGCCCTTTTCGAGACAGGTTCACTCTTTCATGCCCCCTCGAG 7188
Db 7446 TCGACGACGCGCGCAAGAGATCCGAGTTCGATGTGTAATCTCTCAATGCCCTCGAG 7505
Qy 7189 GGGGAGCTTGAAGATCCAGACTGAGGCTGAGCAGGTAGAGCCCAACCCCCCGAG 7248
Db 7506 GAGAGGCAAGGGGAGCCCGACT- - - - - 7528
Qy 7249 GGGGGGCTGGAGCTCCCGGCTCGGACTCGGGGCTGTGCTACTTGTCTCCGAGAGAC 7308
Db 7529 - - - - -CAGCGACGGGTCTTGTCTACTCGGACGGGGAAGCTGT 7568
Qy 7309 GACTCCGTCGTGTGCTGCTCATATGTCACTACTCTGACCGGGGCTTAATACTCTTGT 7368
Db 7569 GAGGAGCTGCTGTGCTCAATGTCTTAATGACAGGTGCTTGAATCAGCCCATGC 7628
Qy 7369 AGTCCGGAAGAGAGATACCGATTAACTCTTGAACAACTCTGTTGGCATATAC 7428
Db 7629 GCTGCGGAGAGAGCAAGTTGCCATCAATCTGTGAGCAACTTGTGCTGCTGACCAAC 7688
Qy 7429 AACAAAGTACTGTACCAACAAGAGGCGCTCACTAAGGGGTAAAGGTAACTTTT 7488
Db 7689 AGTATGTCTACTCCAAATCTCGAGGCAAGTCTGCGGCAAGAAAGGTAACTCTT 7748
Qy 7489 GATAGATGCAAGTGTGACTCTCTACTAGACTCACTGATCTTAAAGACATTTAAGTACG 7548
Db 7749 GACGAGCTGCAAGTCTGAGAGACACTTACCGGAGCTGTCAAGAGATGAAGGAGAG 7808
Qy 7549 GCTCCCAAGTCAACCGCAAGGCTCTCAACATGAGAGAGGTTGCGATTAACTCCACCC 7608
Db 7809 GCGTCCCAAGTTAAGGCTAGGCTTCTATCTAATAGAGAGGCTGCAACATGACCCCA 7868
Qy 7609 CATTTCTCAAGATCTAATAATAGGTTTGGGCTTAAGAGGTCGACAGCTTTCGGGAGG 7668
Db 7869 CATTCGCGCAATCTCAATTTGGCTACGGGCGAAGAGCTCGGAGCTTATCAGCAGG 7928
Qy 7669 GCGGTTAACCAATCAAGTCCGTGTGAGAGACTCTGAGAGACTGAGAAACCAATT 7728

Db 7929 GCGGTCAACCAATCGGCTCCGTGTGGAGGACTTGTGGAAGACCTGAACAACAAT 7988
Qy 7729 CCACAAACCAATTAATGCGCAAAAAATGAGGTGTCTGCGTGGACCCCAACAGGGGCGAAG 7788
Db 7989 GATACCAACCAATCAATGAGAAAAATGAGGTTTTCTGCTCCAAACAGAGAAAGGAGCGC 8048
Qy 7789 AAAGCAGCTGCGCTTAATGTTTTACCTGACCTCGGCTGAGGGGTCTGAGAAATGAGC 7848
Db 8049 AAGCAGCTGCGCTTAATGTTTTACCTGACCTCGGAGGTGAGGTGATGCGAAGATGGCC 8108
Qy 7849 CTTTATGACATTAACAACAAACTTCTCAGCGGTGATGGGAGCTTCTTATGATTCAG 7908
Db 8109 CTTTACAGCGGTGCTCAACCTTCTCAGCGGTGATGGCCCTCATATGAGATTCAG 8168
Qy 7909 TATTCCTCCGCTAGCGGGTGAAGTTCTCTTGAACATGAGGCGAAGAAAGAGACCT 7968
Db 8169 TACTCTCTGCGGACGGGTGAGTTCTGTGTGAATCTGGAATGCAAGAAATGCGCT 8228
Qy 7969 ATGGGTTTTCTGATATACCCGATGCTTGAACCTCAACGCTCACTGAGAGACATCAG 8028
Db 8229 ATGGGCTTCTATATGACACCGCTGCTTGAACCTCAACGCTCACTGAGATGATCCGT 8288
Qy 8029 ACTGAGAGTCCATATATGCGGCTGTCTTTCGCGAGAGGCGCACTGCCATAC 8088
Db 8289 ACTGAGGATCAATTTATCAATGTGTGATGGCCCGCAAGCCAGGAGGCGATAGG 8348
Qy 8089 TCCCTAACTGAGACTTTACGTGGAGGCGCTATGTTCACAGACAAAGGCGCAACTGCG 8148
Db 8349 TCGCTCAAGAGGCGCTTATGTGCGGGGTCCCTGCTAATTCGAAGGGGCAACATGCG 8408
Qy 8149 GGGTACAGGCGTTCGCGCGCGAGCGGGTCTCAACCTAGATGGGAAACACATACAG 8208
Db 8409 GGTATATGCGCGGTGCGCGCAAGTGGGTGCTGACGACTGAGCGCAACCTCTACA 8468
Qy 8209 TGTACTGTGAAGCTTATGCGGCTTGTGAAGCTGCAAGGATTAATGCGCCCAATCTG 8268
Db 8469 TGTACTGTGAAGCTTATGCGGCTTGTGAAGCTGCAAGGATTAATGCGCCCAATCTG 8328
Qy 8269 GTATGCGGCGATGATGTTGTGTCTCAAGAAAGCCAGGGGACCGAGAGAGAGACCG 8328
Db 8529 GTGAACGGAACGACCTTGTGCTATCTGTGAAGTGGGAAACCCAGAGAGATGCGCG 8588
Qy 8329 AACCTGAGCTTCAAGAGGCTATAGCAGGATTTCTGCCCTCTCTGATGACCCCC 8388
Db 8589 GCCCTAGAGCTTCAAGAGGCTATGATGATTTCTGCCCTCCCGGAGACCTGCC 8648
Qy 8389 AGACCGGATGATGATGAGCTGATTAACATCTTGTCTCAATGTGTCTGTGGCGCTG 8448
Db 8649 CAACAGAAATAGCACTTGAAGCTGATTAACATGCTCTCAATGTGTGTGCGGCAC 8708
Qy 8449 GCGCCACAGGCGCGGAGATATCTACTGACCAAGAGACCTTACATCTCCATGCGCGG 8508
Db 8709 GATGCAATCCGCAAAAGGGTGTACTACCTGACCCGTGACCCCAACCCCTCGACGG 8768
Qy 8509 GTCGCTGGGAAACAGTTAGACACTCCCTGTCAATTCATGAGCTGGGAAACATCATCAG 8568
Db 8769 GCTGCGTGGAGAGAGTTAGACACATCTCACTGATCACTCTGCTAGGCAATATATATG 8828
Qy 8629 GCTCAAGACAGCTGACCAAGAACTTGAATGATGATGAGAGGCGGTGTACTCCGTG 8688
Db 8889 GCTCAGAGAGCACTTGAAGAAAGCCCTGATGTGATGATCAAGGCGCTGTACTCAAT 8948
Qy 8689 AGTCCCTTGGACTCCAGCTATATTTGAAGTTACATGGGCTTGAAGCTTTTCTCTG 8748
Db 8949 GAGCAGCTTGAACCTACTCATGATCATTTGAACGATCATGTGTCTTACGCAATTTTCACTC 9008
Qy 8749 CACACATCACTCCCAAGACATGACAGCGGCTGCTTCAAGCTTCAAGAAACTTGGGCGG 8808
Db 9009 CACAGTACTCTCAGGTGAGATCAATAGGCTGCTTCAATGCTCAGAGAAACTTGGGCTA 9068

QY 8809 CCACCCCTCAGAGCGTGAAGAGCCGGGCACTGTCAGTCAAGGCGCTCATCTCCCGT 8868
DB 9069 CCGCTTTGGAGTCTGGAGACATCGGGCCAGAGGTGTCGGCGTAAAGCTACTGTCCAG 9128
QY 8869 GGGGGAGAGCGCGCTTTGGGTGTCATATCTTTCAATTGGCGGTGAAGACCAAGCTC 8928
DB 9129 GGGGGAGAGCGCTGTCATTTGGCGAAGTACCTTTTCAACTGGCGAGTAAAGACCAAGCTT 9188
QY 8929 AAATCTACTCATTTGCGGAGAGCGGCGCTCTGAGATTATCCAGTGGTTTCAACGCTGCGC 8988
DB 9189 AAATCTACTCATTTGCGGAGAGCGGCGCTCTGAGATTATCCAGTGGTTTCAACGCTGCGC 9248
QY 8989 GCCGGCGGGGCGACATTTATCAGAGCGTGTGCGGTGCCGACCCCGCTTATTTGCTCTT 9048
DB 9249 TACAACGGGGAGACATATATCAGAGCGTGTGCGGTGCCGACCCCGTGGTTTATGTG 9308
QY 9049 GCGCTACTCTTACTTTTGTGAGGGGTAGGCTTTTCTTACTCCCGCTCGG 9099
DB 9309 TGCCACTCTTACTTTTGTGAGGGGTAGGCTTACTTACTCCCGCTCGG 9359

RESULT 6

US-09-014-416-6
Sequence 6, Application US/09014416
Patent No. 6153421
GENERAL INFORMATION:
APPLICANT: Yanagi, Masayuki
APPLICANT: Buh, Jeng
APPLICANT: Emerson, Susanne U.
APPLICANT: Purcell, Robert H.
TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
FILE REFERENCE: 20264276
CURRENT APPLICATION NUMBER: US/09/014.416
EARLIER FILING DATE: 1998-01-27
EARLIER APPLICATION NUMBER: US 60/053.062
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 9599
TYPE: DNA
ORGANISM: Hepatitis C virus
US-09-014-416-6

Query Match 45.5%; Score 4136; DB 3; Length 9599;

Best Local Similarity 66.5%; Pred. No. 0; Mismatches 2965; Indels 90; Gaps 5;

Matches 6055; Conservative 0; Mismatches 2965; Indels 90; Gaps 5;

QY 1 ATGAGCAAAATCTTAACTCAAGAAAGAAACCAAAAGAAACCAACCGTCCGCAAA 60
DB 342 ATGAGCAAAATCTTAACTCAAGAAAGAAACCAAAAGAAACCAACCGTCCGCAAG 401
QY 61 GACGTTAAGTTTCGGGCGGCGGCGGCAATCGTTGCGGAGTATCTTTGTCGCGAG 120
DB 402 GACGCAAGTTCCCGGCGGCGTGTGATCGTTGTTGAGTTTACTGTTCCGCGAGG 461
QY 121 GGGCCAGTTGGGTGTGCGGCGGCAAGAAAGTTCGAGCGGTCCGACCACTGTGA 180
DB 462 GGGCCAGTTGGGTGTGCGGCGGCAAGAAAGTTCGAGCGGTCCGACCACTGTGA 521
QY 181 AGGCGCAGCCATCTTAAAGATCGGCGCTCACTGCGCAATCTTGGGGAAGAAACAGAA 240
DB 522 AGGCGCAACCTATCCCAAGGCTGCGGACCGAGGGGAGGGCTTGGGCTCAGCGCGG 581
QY 241 TACCCCTGCGCTTATACGGAATGAGGACTCGGCTGGGCGAGATGCTCTGTCGCC 300
DB 582 TACCTTGGCGCTTATGCAATGAGGCTGCGGCTGGGCGAGATGCTCTGTCGCC 641
QY 301 CGAGTTTCCGCTCTTTGGGCGCCCAATGACCCCGCGCATAGTCCGCGCAAGTGGT 360
DB 642 CGGCGTCCCGCTTATGTTGGGCGCCCAAGGAGTCCCGCGCATAGTCCGCGCAAGTGGT 701

QY 361 AAGTTCATGATACCTTAACTGTCGCTTTGGCGAAGCTTATGAGGTATCTCCGTGTG 420
DB 702 AAGTTCATGATACCTTAACTGTCGCTTTGGCGAAGCTTATGAGGTATCTCCGTGTG 761
QY 421 GGGCGCCCGCTGGCGGCGCTGCGAGAGCTTCGCGCATGAGGTATCTTGAAGAGC 480
DB 762 GGGCGCCCGCTGGCGGCGCTGCGAGAGCTTCGCGCATGAGGTATCTTGAAGAGC 821
QY 481 GGGGTAATTTTGAACAGAGAACTTACCGGCTGCTTTTCTTATCTTTGCTGAGC 540
DB 822 GGGGTAATTTTGAACAGAGAACTTACCGGCTGCTTTTCTTATCTTTGCTGAGC 881
QY 541 CTGCTGTCTGATCAACCAACCCCGTCTGCGTCCGGAAGTGAAGAAATCACTACGCGC 600
DB 882 CTGCTGTCTGATCAACCAACCCCGTCTGCGTCCGGAAGTGAAGAAATCACTACGCGC 941
QY 601 TACATGATGATACAGATGACCAATGACAGATTACTGGCAGCTCCAGGCTGTGTC 660
DB 942 TACATGATGATACAGATGACCAATGACAGATTACTGGCAGCTCCAGGCTGTGTC 1001
QY 661 CTCACGTCCCGGAGTCCCGTCCGGAAGAAAGTGGAGAAATGATCTGAGTCTGAT 720
DB 1002 ATGCAATATCTCCGAGTCCGCTGCTGCTTTCAGAGAGGTAAAGCTCCGTTGCTGAGT 1061
QY 721 CCGGTCTACCGAATGTGCGCGTGCAGCGGCGCGGCGCTGACCGAGGCTTGGCGAGC 780
DB 1062 GCGCTACCTCCACGCTCGCGGCAAGAAATGCGAGGCTCCGCTACGAAATACGACGC 1121
QY 781 CACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 1122 CAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1181
QY 841 TGGCGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 1182 TGGCGATCTATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1241
QY 901 GTCCAGATCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
DB 1242 GTCCAGATCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1301
QY 961 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 1302 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1361
QY 1021 CCGGAGTCTTATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1362 CCGGAGTCTTATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1421
QY 1081 TACTTCTTATGACAGGAGCGTGGGGAAGTGTGATCTTGTGCGCGCGG 1140
DB 1422 TACTTCTTATGACAGGAGCGTGGGGAAGTGTGATCTTGTGCGCGCGG 1481
QY 1141 GTGAGCGCGGCACTATCTGTTGGGCTTCTGCGCGGCAAGCAACCGGCGCTCAC 1200
DB 1482 GTGAGCGCGGCACTATCTGTTGGGCTTCTGCGCGGCAAGCAACCGGCGCTCAC 1541
QY 1201 AGCTTATTTGACATGAGGCGCCAGGCAAGAAATCCAGTCTGTTTAAACCAATGAGCAGTGG 1260
DB 1542 TCCCTTTTCTATCTGAGGCGCTCTCAGAAATCAAGTGTGTAATACCAAGGCAAGTGG 1601
QY 1261 CAGATCAACCGCAACCGCGCTGAACTGCAATGACTCTTGGACACCGGCTTTATCGGCTT 1320
DB 1602 CAGATCAACCGCAACCGCGCTGAACTGCAATGACTCTTGGACACCGGCTTTATCGGCTT 1661
QY 1321 CTGTTTCAACCGCAACCGCTTCACTGTCAGAGATGTCGGAAGCATGTCGCTGCGCG 1380
DB 1662 CTGTTTCAACCGCAACCGCTTCACTGTCAGAGATGTCGGAAGCATGTCGCTGCGCG 1721
QY 1381 AGTATGAGGCTTCCGAGTGGATGAGGCGCTTCAATATGAGATATGTCACCAAT 1440
DB 1722 CCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1780

OY	1441	CGAAGAGATATGAGACCTTATGTGCTGGACACTAACCCACCAAGCAGTGTGGGTGTCTCC	1500
Db	1781	-----GGATCAGAGAGCCTTATTTGGCTGGCAATTACGGGCTCCACGGTGTGGTGTCAACC	1835
OY	1501	GCGAAGATCTGTGTGTGGGCCAGTGTACTGTTTCAACCCCAAGCCAGTGTGTGGGACG	1560
Db	1836	GCTTCGACAGTGTGTGTGTGCATGTATTTGTTTCAACCCCAAGCCCTGTTGTGTGTGGAC	1895
OY	1561	ACCGACAGGCTTTGGAGCGCCCACTTACACGTGGGGGGAATGAGACAGATGTCTTCCTA	1620
Db	1896	ACCCATCGTTCCGGTGTCCCTTAACGTATAGCTGGGGGAGAAATGAGACAGATGTATCTC	1955
OY	1621	TTGAACAGCACTCCACCAACCGCTGGGGTCAATGTTCGGCTGCACGTGGAATGAATCTTCT	1680
Db	1956	CTCAACCAACACGCGTCCGCAACAGGCAACGTGTTCGGCTGTATACATGGAATGAATGACT	2015
OY	1681	GGCTATACCAAGACTTGTGGGGGACACACCTGTCCGCTATAGAGCTTACAGCCGACG	1740
Db	2016	GGGTCTACTTAAGACTGTGGGAGATCCCCGTGTATCATCGGGGGGTGTGTAACCGAC	2075
OY	1741	ACGGAAGCTGTTGTGCCCAACGGAAGTGTTTTGAAGAGCATCTGTATACCACTTACCTCAA	1800
Db	2076	T-----TGATCTGGCCACGAGGACTGTTCCGGAAGACCCCGAGGCTACTTACACAAA	2129
OY	1801	TGCGGCTCTGGGGCCCTGTGCTACGCGCAAGTGTGCTATCGACTACCTTACAGGCTGTGG	1860
Db	2130	TGTGGCTCTGGGGCCCTGTGTGACACTGTAGGTCTAGTATGACTTACCACTACAGGCTTTGG	2189
OY	1861	CATTACCCCTGCACAGTTAATTAATCACTTCAAAATTAAGAGATATGTGGAGGGGTT	1920
Db	2190	CATTAACCTCTGCACCTTCAATTTTTCATCTTTAAGTTAAGTAAGTATGTGGGGGGGTG	2249
OY	1921	GAGCACAAGGCTCAGGGCTGCATGTGCAATTTCACTGTGGGGAATCGTTGCACTTGAAGAC	1980
Db	2250	GAGCACAAGGCTCAATGCCGATGCAATTTGAAGCTGAGAGAGAGCCCTGTATTTGAGAGAC	2309
OY	1981	AGAGACAGAGTCAACTGTCTCTCTTTGTGCACTTCACACGAGATGGGCCATTTTACT	2040
Db	2310	AGGATATAGGTACAGAACTCAGCCCGCTGTGCTGTCTACACAGAGTGGCAGATCTCGCC	2369
OY	2041	TGCTCTTAATCTGGAGCTGCCCGCTGTGTGCATGTGTCTTCTCCACTTCACCAAAATATC	2100
Db	2370	TGTCTTTTCAACACCTTACCGGCTTTTACCATGTGGTGAATTCATCTTCATCAACAAATC	2429
OY	2101	GTGACGCTACATTCATGTATGTGGCTTATCACTGTGCCCTCAACAAATATCATGTCCGATGG	2160
Db	2430	GTGACGCTGCAATACCTGTACGGGTGTAGGGTCAAGCTTTGTCTCTTTGCAATCAAAATGG	2489
OY	2161	GAGTGGGTAACTCTTAATTCCTGTCTCTTACGGAAGCGCAGAGTTTGGCGCTGTATGG	2220
Db	2490	GAGTAAATCTGTGTGCTTTTCTTCTCTCTGACAGCGCGCGTGTGTGCTGTGTGG	2549
OY	2221	ATGCTCATCTTGTGTGGGCGAGGCGAAGACACTAGAGAAGCTGTTCATTTGCAGGCT	2280
Db	2550	ATGATGTGTGCAATAGCCACAGGCTGAGGCGCGCTTGAAGAACTTGTGTGTCTCAAAATGG	2609
OY	2281	GCGAGCGGAGTACGTGCAATGGGCTCCCTAATTTTGTCACTTTTTCGTGGGCTTTGG	2340
Db	2610	GCGTCCGTGGCGGAGCGCATGTATTTCTTCTTCTTGTGTCTTGTGCGCGGCTGG	2669
OY	2341	TACATCAAGGGTGGGATAGTCCCTGTAGCTACCTATTTCCCTCACGTGGCTGTGTCTTT	2400
Db	2670	TACATTAAGGGCAGAGGTGGCTCTGTGGGGCGGCGTATGTTTATATGGGTATGGCCGCTG	2729
OY	2401	AGCTTACTGTCTCTACGACTTGTGCCCAACAGGCTTATGTCTTATGACGATCTGTGCATGGC	2460
Db	2730	CTCTGTCTCTTACGTGGCGTTTACACACAGACATATGACATGACACAGAGGTGGCGCG	2789
OY	2461	CAGATATGAGAGGGGCTGTGCTGTATATATATCATCTCTTATATCTACACCCCGGGATATAG	2520
Db	2790	TGTGTGTGGGCGGCTTTGTTCTTGTGGGTTAATGGCGCTGATCTGTGTGCCATATTACAG	2849
OY	2521	ACCTTTCTCAGCCGGTTTTTGTGTGTGTGTGTCTATCTTGCACCTGGGGGAAGCTATG	2580

Db	2850	CGTATATCACTGCTGTCATGTGGTGGCTTTCAGTATTTTCTGACCAAGTAAAGGCCAA	2909
Oy	2581	GTCCAGAGTGGGACCAACCTATGCAAGGTGCGCGGTGACCGTGATGTGATCATATGTGGCC	2640
Db	2910	CTGCACGTGTGGTGTTCCTCCCTCCAAAGTCCGGGGGGGGCGGATGCGTCATCTTATCTC	2969
Oy	2641	GTCCGCATATTCTAACCAAGGTGTGTGTGACATAACCAAGTGGCTTTGGCGGTCTT	2700
Db	2970	ATGTGTGTAGTACACCCGACCCCTGTATTGTGAACATACCAAACTATCTGGCCATCTTC	3029
Oy	2701	GGGCTGTGTTACCTCCATAAAGTGTGTTGACGGCGGTGTCCTTGTCAGGGCTCAC	2760
Db	3030	GGACCCCTTTGGAATTTCTCAAGCCAGTTGTGTTAAAGTCCCTTACTTGTGTGGTTCAA	3089
Oy	2761	GCTCTACTGAGAGATGTGCACACCATGGCAAGGATCTCCGGGGGGCAGGTACGTCCAGATG	2820
Db	3090	GGCTTCTCCGGATCTGTGCGGTGACCCGGGAAGATACCCGGAAGTCATTAAGTCCAAATG	3149
Oy	2881	GCGCTACTAGCCCTTGGCAAGTGGACTGGACCTTATCTATGACCACTTCACTCCCTATG	2880
Db	3150	GCCATCTCATCAAGTTAGGGGCGCTTACTGGCACTATGTATTAACATCTCACCCCTCTT	3209
Oy	2941	AGTCCGATGTGAGAAAGATCATTTGTGTGGGAGACGGACAGCTGTTGGGGGACATT	3000
Db	3270	TCCCGAATGGAGACCAAGCTCATACGTGGGGGGGAGATACCGCGGTGCGGTGACATC	3329
Oy	3001	TTACACGAGCTTCCCGTGTCCGCCGACTTGTGTGGAGGTCTCTTGTGCCACGTGAT	3060
Db	3330	ATCAACGGCTTGGCCGCTCTGTGCCCGTAGGGGCGAGGAAGATACGTCTTGGGCGACGGAC	3389
Oy	3061	GGCTATACCTTCAGAGGGGTGGAATCTTGTGCCCCCATCACTGTTACGCCCAAGACACA	3120
Db	3390	GGAATGTGTCTCAAGGGGTGAGGATCTGTGGCGCCATCAAGCGGTACGCCCAACAGCG	3449
Oy	3121	CGTGGCCCTTTGGGACACCATATGTGTGAGCATGACGGGGGGCGACAGAACAGAGCT	3180
Db	3450	AGAGGCTCTCTAGGGGTGATATATACAGAGCTGTAGCTGGCGGACAAAACCAAGTGGAG	3509
Oy	3181	GGGGAATTCAGGTCTGTGTCCACAGTCACTCAGTCTTCTCCGGAACATCATCTCGGG	3240
Db	3510	GGTAGAGTCCAGATTCGTGTCAACGTCAACCAACCTTCTCGGAGAGTGCATCAATGGG	3569
Oy	3241	GTTTGTGAGCTGTCTACCATGAGCTGGCAACAAGCTCTGGCGGCTCACGGGGTCCG	3300
Db	3570	GTATGCTGGACTGTCTACCAAGGGGCGGAAACGAGGACCAATCGCATCAACCAAGGCTCT	3629
Oy	3301	GTCACGAGATGTATCTCAGTGTCTGAGGAGGGGACCTTATAGGGTGGGCCACGCCCTCGGG	3360
Db	3630	GTCAATCAAGATGTATCAATGTGGACCAAGACTTGTGGGCTGGCCCCCTCTCAAGGT	3689
Oy	3361	ACTTAATCTTTGAGCGCGTGCACAGTGTGAGACGGTGCACCTGTATCCTGTACCGCGGAC	3420
Db	3690	TCCCGCTCATTTGAACCTGTACTGTGGGCTCTCCGGACTTTTACTCGTGTCAAGAGCAC	3749
Oy	3421	GCTGATGTCAATCCCGGCTCGAAGACGCGGGGACAAACGGGGAGCGCTACTCTCCCGAGA	3480
Db	3750	GCCGATGTCAATCCCGTGGCGGGCGAGGTGATGACAGGGGTAGCTGTTTGGCCCCGG	3809
Oy	3481	CTCTTTTCAACCTTGAAGGGGTCTCTCAGAGAGCCCGGTGTATGCCCCAGAGGGCAAGCT	3540
Db	3810	CCCATTTTCTAATTGAAGGCTCTCTCGGGGAGTCCGCTGTGTGTGCCCGCGGACAGAGCC	3869
Oy	3541	GTCCGAGCTTTTCCGGGACGCTGTGTGTCTCTCGGGGCGCTGAGTCAATAGATTTTCATC	3600
Db	3870	GTGGGCTTATTCAGAGGCGCGGTGTGTGACCCGTGTGAGTGGCTAAAGCGGTGACTTTATC	3929
Oy	3601	CCGTTGAGACCTCGACATGTGTACGCGGTCCCCCACTTTATGTGACCAACAGCACCA	3660

Db 3930 CCTGTGGAGAACTTAGGGACAACCATAGATCCCCGGTGTTCACGACAACTCCCTCCCA 3989
Oy 3661 CCTGTGTGCCCCAGACCTATCAGGTGG99TACTTGTGATGCCCCGAGCTGGAGTGAAG 3720
Db 3990 CCAGCATGCCCCAGAGCTTCCAGGTGGCCACCTGCACTGCTCCACCGGACGGGTAG 4049
Oy 3721 AGCACCAGTTCCTGTGTCATATGCTGCTCAGGGGTATMAAGTCTAAGTCTTATCC 3780
Db 4050 AGCACCAGTTCCTGCGTGTGTAAGCAGCCAGGGCTACMAAGTGTGTGCTCAACCCC 4109
Oy 3781 TCAGTGTGTCACCTCGGGTGTGGGGCTGACTTGTGTAAAGCAGATGTCATATCC 3840
Db 4110 TCTGTGTGCAAGCTGGGCTTTGTGTCTTACATGTCCAAAGGCCATGGGGTGTGATCT 4169
Oy 3841 AACATTAGACTGTGAGTCAAGACTGTGACGACCGGGGGCCCATCAGTACTCCACATAT 3900
Db 4170 AATTATCAGACCGGGGTGAGAACTTACACTGGCAGCCCCCATCAGTACTCCACTAC 4229
Oy 3901 GGCAAATTCCTGCGCGATGGGGGCTGTGCGGGCGGCGCTTACGACATCATATGTGAT 3960
Db 4230 GGCAGATTCTTGGCCGACGGCGGTGCTCAGAGAGTGTCTTATGACATMAATTTGTGAC 4289
Oy 3961 GAATGCCATGCGGTGAGTCTTACCACTATCCTTGGCATCGGACAGTCTTATCAAGCA 4020
Db 4290 GAGTGCCACTCAGCGATGCCACATCTTGGGCATCGGCACTGTCTTGAACMAAGCA 4349
Oy 4021 GAGCAGCTGGGGTCAAGCTAAGTGTGCTACAGCTACGCCCCCTGGGTCAGTACA 4080
Db 4350 GAGATGCGGGGGCGAGACTGTGTGTGCTGCGCACTGTACCTCCGGGTCCGTCACT 4409
Oy 4081 ACCCCCAACCCCAATAGAGAGGTGGCCCTTGGGAGAGAGGCGAGATCCCTTCTAT 4140
Db 4410 GTGTCCCATCTCAACATGAGAGAGTTGCTGTTCACACCGGAGAGATCCCTTTTAC 4469
Oy 4141 GGGAGGGGATTTCCCTGTCTTATCATCAAGGAGAGAAACATCTGATCTTGTGCATTTCA 4200
Db 4470 GGCAGAGCTATCCCTCTGAGGTGATCAAGGGGAGAGACATCTATCTTGTGCACCTCA 4529
Oy 4201 AAGAAAAAGTGTGACGAGCTGCGGGCGGCCCTTGGGGGTATGGCTTGAATCTAGTGGCA 4260
Db 4530 AAGAAAGTGTGACGAGCTGCGGGCGGCCGAGCTGTGATGGGATCAATGCGGTGGCC 4589
Oy 4261 TACTACAGAGGGTGTGAGCTGTCCGTATACCACTCAGGAGAGAGTGTGTGCGCC 4320
Db 4590 TACTACCGCGGTCTTGAAGTGTGTGTATCCGACACGAGCGGAGTGTGTGTGTGCG 4649
Oy 4321 ACCGACGCTTCATACAGGGTATACCTGGGACTTTGATCTCGGTATGTGATCTGCAAGTA 4380
Db 4650 ACCGATGCTTCATACAGGGTATACCTGGGACTTTGATCTCGGTATGTGATCTGCAAGTA 4709
Oy 4381 GCGGTCACTCAAGTTGTAGACTTTCAGTTTGAACCCACATTCACCATMAACACAGATT 4440
Db 4710 TGTGTCACTCAGACAGTGTGATTTGACCTTGAACCTTACCTTTACCTTGAAGACACACG 4769
Oy 4441 GTCCCTCAAGACGCTGTCTCAGTAGAGCCAGCGCGGGGTGACAGGGTGAAGGAAAGTG 4500
Db 4770 CTCCCCAGAGTGTGTCTCCAGGACTCAACGCGGGGAGAGACTGGAGGGGAAAGCA 4829
Oy 4501 GGCATTTTATAGTATGTTTCCACTGTGTGAGAGGCTCAGAGAAATTTGAAGTGTAGTG 4560
Db 4830 GGCATCTATAGATTTGTGGCACCGGGGAGCGCCCTCCGGCATGTGTGACTGTGCTGCTC 4889
Oy 4561 CTCTGTAGTGTCTACGACGAGGGGCGCATGTGTAGACTCACACATCGGAGCCACC 4620
Db 4890 CTCTGTAGTGTCTACGACGCGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4949
Oy 4621 GTGAGGCTCAAGGCGTATTTCAACAGCGCCGGTGTGCTGTGTGTGTGTGTGTGTGTGTGT 4680
Db 4950 GTTGGGCTACGAGCTATACGAGCAACCGGGGCTTCCGTTGTGTGTGTGTGTGTGTGTGT 5009
Oy 4681 TTTTGGGAGGAGTTTTCACCGGCTTCAACACATAGATGCCCACTTCTTCCCAACA 4740
Db 5010 TTTTGGGAGGAGGTTTTCACCGGCTTCACTCATATAGATGCCCACTTCTTTCACAGACA 5069

Oy 4741 AAGCAATCGGGGAGAAATTTTCGACTACTTAACAGCTTACAGGCTACAGTGTGCGCTAGG 4800
Db 5070 AAGCAAGTGTGGAGAACTTTCTTACTGTGTAGCTTACCAAGCCACGTTGTGCTTAGG 5129
Oy 4801 GCCAAAGCCCCCGCTGTCTGTGGAGCTCATGTGTGAAGTGTGTGACTGTCAAGCCC 4860
Db 5130 GCTCAAGCCCCCTCCCATCTGTGTGGACAGATGTGTGAAGTGTGTGTGTGTGTGTGTGTGT 5189
Oy 4861 ACATGTGTGGGCCCACTCTCCCTGTACCGCTTGGGCTCTGTTAACCAAGAGTCAACC 4920
Db 5190 ACCCTCATGTGGCAACACCCCTGTGTACGACTGTGGGCGTGTGTGTGTGTGTGTGTGTGT 5249
Oy 4921 CTCACATTCGCGTGAAGAAATACATGTGCACCTGTGATCAAGCCGCTTGAAGTCAATG 4980
Db 5250 CTGACGCAACCAATACCAAAATACATGACATGTGATGTGCGCGGACCTGTGAGGTGTGTC 5309
Oy 4981 ACCAGCAATGGGTCTTGTGGCAGGGGAGTCTTGGCGGCGCTGCGCGCTATTTGCTGGG 5040
Db 5310 ACCAGCACTGGGTGCTGTGTGTGGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCA 5369
Oy 5041 ACCGGGTGTGTGTGATCATGTGCGGCTTGCATTTAACAGGAGCGCGTGTGCGCG 5100
Db 5370 ACAGGCTGT 5429
Oy 5101 GACAAAGAGTCTCTATGAGGCTTTGTATGATGAGATGAGAAATGTGCTCTAGGGCGGCT 5160
Db 5430 GACAGGAGGTTCTTACAGAGAGTTCATGTAGATGTGAAGAGTGTCTCAGGACTTACCG 5489
Oy 5161 CTCATTGAAGAGGGGAGCGGATAGCCGAGATGTCTGATCCAAATCCAGGCTTATTTG 5220
Db 5490 TACATGAGCAAGGATGT 5549
Oy 5221 CAGCAAGCTTCCAAAGCTCAAGCATTAACCACTGTGAGGCTTCAATGGCCCAAG 5280
Db 5550 CAGACGCGTCCGCGCATGAGAGTGTATCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5609
Oy 5281 GTAGAACAATCTGTGGCCAAACACATGTGTGAATCTTATTAAGCGGACTTCACTGCA 5340
Db 5610 CTGAGGTCTTTTGTGGGAGAGCAATGTGGAATTTATCATGATGAGATCAATATCTTGGG 5669
Oy 5341 GAGCTATCAACACTGCGAGGAAACCTGCAAGTGTCTTCAATGTGCGCTTCAAGTCCGCG 5400
Db 5670 GGCCTGTCAACGCTGT 5729
Oy 5401 CTCACAGTCCGCTGTCAACAGCACACATATCTTCAATTTTGGGGGGCTGAGCTTA 5460
Db 5730 GTACCAAGCCCACTTAACACTGTGACCAACCTTCTTCAACATATTTGGGGGGGTGGTG 5789
Oy 5461 GCATCCCAATTTGCACACACCGCGGGGCGCACTGTGCTGTGTGTGTGTGTGTGTGTGTGT 5520
Db 5790 GCTGCCAGTGTGCGCGCGCGCGCGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5849
Oy 5521 GCTGCCGTAGGCGATATAGCTTATAGTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5580
Db 5850 GCCGCATTCGCGAGCGTGTGACTGTGGGAAAGTGTCTGTGTGTGTGTGTGTGTGTGTGT 5909
Oy 5581 GCGGGGCAATTTGGGGGGCTGT 5640
Db 5910 GCGGGGCTGTGCGGAGGCTTGT 5969
Oy 5641 GAGAGT 5700
Db 5970 GAGGACTGT 6029
Oy 5701 ATCTGTGCGGCGCATTTGTGCGCGGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5760
Db 6030 GTCTGTGCGCAATATCTGCGCGCGGCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6089
Oy 5761 AATGACTCATGTGCTTGT 5820
Db 6090 AACGGTGTATGT 6149

QY	5821	GAGTCGGA	TGGGTGCGACGCGGTGACCCAA	CTACTTGGCTCCCTTACATTAACAGGCTG	5880	
Db	6150	GAGGCGGATG	CAGCGCCGCCGCTCACTGGCACTAC	CAGACGCTCACTGTAAACCAGTCC	6209	
QY	5881	CTCAGAAGACT	CCAACTGGATTACTGAGGACTGCCCAT	CCCATCGGCGGCTCTGTGG	5940	
Db	6210	CTGAGGGGCA	CTGCATTCAGTGGATTAAAGCTCCGAGTGTAC	CACTCATGTCTCGGTTCTCGG	6269	
QY	5941	CTCCGCGAT	TGTTGGGACTGGGTTTGACCACTCTTA	ACAAGCTTTAAATAATGGCTGACC	6000	
Db	6270	CTAAGGGGCA	CTGTGGGACTGTGATGTGCGAGGTGTGAG	CGCACTTTAAGACCTGTGCTAAA	6329	
QY	6001	TCCAAATTA	TTCGAAAGA	TGCCCCGCTCCCTTTGTCTCTGTCA	AAAAGGGGTTACAG	6060
Db	6330	GCCAAAGCT	CAATGCCAACACTGCTCTGGGATTCCTCTT	TGTCTCTCCACAGCGCGGGTAAATG	6389	
QY	6061	GGCGTGTG	GGCGCGGCACTGGCATGTATGCAACACAG	GTGTCCTTGCGCGCCCAATATCTCT	6120	
Db	6390	GGGGTCTGT	GGAGAGACGGCATTTATGCAACCTGCT	CCACTGTGGAGCTGTGAATCACT	6449	
QY	6121	GGCAATGT	CCGCTTGGGCTCCATGAGATCACGGG	GGCTTAAGACCTGATGATATCTGG	6180	
Db	6450	GGACATGT	CAAAAACGGGACGATAGAGATCGT	CGATCTTAGACCTGCAAGGAATGTGG	6509	
QY	6181	CAGGGGAC	CTTTTCTATCAATTTGTTACAGGAGGG	CGGACGTGCGCGAAACCCGCGCCA	6240	
Db	6510	AGTGGGAC	GTTCCTCCCATTTAAGCCCTTACACAC	CGGGCCCTGTGTACTCTCCCTTTCGCGCG	6569	
QY	6241	AACTTTAA	GGTGGCATCTGTGAGGGGTGGCGGCT	CAGAGTACGGGAGGTGACGCAAC	6300	
Db	6570	AACCTATAG	TTTCCGGCTGTGAGGGGTGTCTG	CAGAGATATCGTGGAGATTAAGCGGGGTG	6629	
QY	6301	GGGTCA	TACCACTAACAGAGCTCACCACTGAT	AATCTTGAAGTCCCTGTGCCAATA	6360	
Db	6630	GGGGAC	TTTCCATAGTACGGGATATGCTAC	TGCAATCTTAAATGCCGTGCAAGTCC	6689	
QY	6361	CCCTCT	CCCGAGTTCTTTTCTGGGTGAGCG	AGTGTCAATAGTTTGGCCCCCA	6420	
Db	6690	CCATGCGC	CGCAATTTTTCACAGAAATTTGG	ACGGGGTGGCTTACACAGTTTGGCCCCCT	6749	
QY	6421	CCGAAGC	CGTTTTTCCGGAGTAGAGTCTCGT	CTGCGTGGGCTTAAATTAATTTGTG	6480	
Db	6750	TGCAAGC	CTTGTCTGCGGAGAGGATATCAAT	TACAGATGAGACTTCCACAGATAC	6809	
QY	6481	GGGTCC	AGCTCTCTTGGGACCTTGAACCCG	ACACAGACGATATGATGCTCATTA	6540	
Db	6810	GGGTGCGA	TTTAACTTGTGGACCCGAAACG	SACTAGCGGTGTGAGTCCATCTCACT	6869	
QY	6541	GATCAT	CTCATATCACGGCGGAGACTTG	CAGCGCGGCGTATTAGCGCGGGGTCA	6600	
Db	6870	GATCCCT	CCCCCATTAATPACAGCAGAGG	CGGCGGGAGAAAGTTGGGAGAGGGGTCA	6929	
QY	6601	TCCGAG	CGAAGTCTCTGGGCGAGGCACTAT	TCGGCAACATGCGTGGCAGCACTTG	6660	
Db	6930	TCTATG	CGACGCTCTCGGCTTAGCGCAG	GTGTCGCTCATCTTCCAAAGGCACTTG	6989	
QY	6661	ACCACG	CGCAAGCTATGATGTGACATG	GTGATCTAACT-----GTTC	6708	
Db	6990	GCCAA	CACTGACTCCCTGACGCGGAGCT	CTATAGGCTAACTCTCTGTGAGGCA	7049	
QY	6709	ATGGG	GGGCGATGTGACTCGGATAG	AGTCTGGGATCCAAAGTGTGTTCTG	7068	
Db	7050	ATGGCG	CGGCAATCAACAGGGGTG	AGTACAGAGAAAGGTGTGATTTCTG	7109	
QY	6769	GACCA	ATGTGTGAAAGAGAGGACCTT	GAAGCTTCGATACATACAGAAATTA	7169	
Db	7110	GATCC	GCTGTGTGGCAGAGAGATG	AGGAGGATCTTCGTAACCTGCA	7199	
QY	6829	CCCAAG	AGAGGTTTCCACACAGCTT	TACCGGCTGTGGGACGCGCTG	7249	
Db	7170	AAGTCT	CGGAAATTTGCGCCGGG	CCCTGCGCTGTGGGCGGGCGG	7299	
QY	6889	CTTGTG	ATGTGTGAAAAGGCCAGATTA	CCACCGGCACTGTTGCGGCTGTG	6948	

Db	7230	CTAGTTAGAGAGCTGGAAAAAGCCCTGACTAGAACACACTGTGTGCTCATGTGGCTGCCCTTA	7289
Qy	6949	CTTCCTCTTAGAAAAACCCCGAGCGCTTCCCCCAAGAGCGCGGAGACATGTGGGCTTAAGT	7008
Db	7290	CCACTCTCAACGGTCCCTCTCTGTGCTCTCCGCTCGGAAAAAGAGCTACGGTGTCTCAAC	7349
Qy	7009	GAGGACTCTCATAGAGATGCCCCCTTCAACAGCTGGGCCATTAAAGTCTTTTGGCAGACCCCC	7068
Db	7350	GAATCAACCCCTATCTACTGCTTGGCCGAGCTTGCCACAAAAGTTTGGAGCTCTCTCA	7409
Qy	7069	CCAAGCGGCAATTACAGGCTTTTCCACGGGGGCGGCGCTGCGCATTCGCGAGCTCAGACG	7128
Db	7410	ACTTCGCGGCATTAAGGGCGACAAATACACAACTCTCTGTAGCCCCGCT-----	7457
Qy	7129	CCTTCCTGATGATTTGAGCCCTTTTGGAGACAGGTTCCATCTCTTCATGTCCCCCTTCGAG	7188
Db	7458	CCTTCTGTGCTGCCCCCCCCGACTCCGACGTTGAGTCTATTCTTCATATGCCCTCCCTGGAG	7517
Qy	7189	GGGGAGCTTTGGAGATCCAGACTGTGAGGCTGAGCAGGTAGAGCCCCCAACCCCCCCAG	7248
Db	7518	GGGAGGCTGTGGGATCCGGA-----	7537
Qy	7249	GGGGGGGTGGAGGCTCCCGGCTGGACTCGGGGTCTGTCTATTTGCTCCGAGAGGAC	7308
Db	7538	-----TCTGAGGAGCGGGTCAATGTGACGGTCACTAGTGTGGGCGCGACACG	7583
Qy	7309	GACTCCGCTGTGTGCTGCTCATGTCAATCTCTCGACCGGGGCTCTTAATTACTCTTGT	7368
Db	7584	GAAGATGTCTGTGTCTCTCAATGTCTTAATTCGTGACAGCGGACATCTGACCCCGTGC	7643
Qy	7369	AGTCCGAGAGAGAAAGTTAACGATTAACCCCTTGAGCACTCTCCGTGTGGATATCAC	7428
Db	7644	GCTCGGAGAGAAACAAAACCTGCCCATCAAGCACTGAGCAACTGTGTGTCAACGCCATCAC	7703
Qy	7429	AACAAGGTGATCTGTACCAACAAGAGGCGCTCATAGAGGCTTAAGAGCTTAATCTTTT	7488
Db	7704	AATCTGTATGTATCTACACACTTCAACGCAATGCTTGCACAAAGGCAGAAAGAAAGTCAATTT	7763
Qy	7489	GATAGAGTGCAGAGTCTCGACTCCCTTACTAGACTCACTCTTAAAGACATTTAAGCTAGCG	7548
Db	7764	GACGAGACTGCAGATTTCTGGACAGCATTAACAGAGCTGTCTCAAGAGAGTCAAGAGCAGCG	7823
Qy	7549	GCTCTCAAGGTCAACCGAGAGCTTCTCAGCATGAGGAGGCTTGCAGTTAACCCACCC	7608
Db	7824	GCGTCAAAAGTGAAAGGCTTAACCTGTCTATGCGTAGAGAAACCTTGCACTGACGCCCCA	7883
Qy	7609	CATTCTGCAAGATTTAATATGGGTTTGGGGCTTAAGAGGTCCGCACTTGTCCGGGAGG	7668
Db	7884	CATTCAACCCAAATCCAAAGTTTGGCTATGGGGCAAAAACATGCTCGTTTGCATATGCCAAGAG	7943
Qy	7669	GCGGTTAACCAACATCAAGTCCGTGTGAGAGGAACTCTCGAGGACTCAGAAAACCAATTT	7728
Db	7944	GCGGTAGCCCAATCACTCCGTGTGAGAAAGACCTTTCTGGAGAGCAGTGTATACCAATTA	8003
Qy	7729	CCCAACAACATTAATGCCCCAAAAATGAGGTGTTCTGCGTGAACCCCAACAAGGGGGCAAG	7788
Db	8004	GACACTACATCAATGAGCCCAAGAAAGAGGTTTCTGTGCGTTGACGCTGAGAAAGGGGGGTGCT	8063
Qy	7789	AAAGCAGCTGGCTTATGTGTTTACCTTGACTCTGGGCTCAGGGTCTGCGAGAAATGTGGC	7848
Db	8064	AAAGCAGCTGTCTCAATCGTGTGTCCCGACCTGGGCGTGGCGCGTGTGCGAAMAATGGCC	8123
Qy	7849	CTTTATATGACATTAACA AAAAATCTCTCAGGCGGTGATGAGGGGCTTCTTATGATTTCAAG	7908
Db	8124	CTGTACGACGTGTGTAGCAAGCTCCCTCTGGCGGTGATGGGAAGCTCTACGATTTCCA	8183
Qy	7909	TATTCGCCCGCTCAGCGGGTAGAGTTTCTCTTGAAGCATGTGGCGGAAAAAGAGACCTT	7968
Db	8184	TACTACCAAGAGACGCGGGTTGAATTTCTGTGTCAACGTGTGAAGTCCAAGAGAACCCCG	8243
Qy	7969	ATGGGTTTTTGTATGATACCCGATGTTTGACTCAACCGTCACTGAGAGAGATACAGG	8028

Db	8244	ATGGGGTTCTGTGTGATGATACCCGGCTTTTGATCTCAACAGTCACTGAGAGGACATCCGT	8303
Oy	8029	ACTGAGAGTCCATATATATCGGGCCCTGCTCTCTTGGCCGAGAGGCCCACTGCTCATAC	8088
Db	8304	ACGAGAGGAGAAATTTATCCAAATGTTGTGACCTTGACCCCAAGCCCGGTGGCATAAAG	8363
Oy	8089	TCGCTAACTGAGAGACTTTACGTGGAGGGGCTATGTTCAACAGCAAGGGGCAAACTGC	8148
Db	8364	TCCTCACTGAGAAGCTTTATGTTGGGGGCCCTTTACCAATTCAAGGGGGGAAAATCTCG	8423
Oy	8149	GGGTACAGGCGTTGCCCCGCCAGCGGGGTGCTCACTACATGGGGAAACCATCA	8208
Db	8424	GGCTAACCGCAGGTGCGCGGCGAGCGGCTACTGCAAACTACTGTGTAAACCTCTACT	8483
Oy	8209	TGCTAACGGAAGCCTTAGCGGCTTTAACTGCAAGGATATATCGGCCCAATGCTG	8268
Db	8484	TGCTACATCAAGGCCCGGCGAGCCTGTGAGACCGGAGCTTCAGAGACTGCACATGCTC	8543
Oy	8269	GTATGCGGCGATGACTGTGTTGTATCTCAAGAAAGCCAGGGGACCGAGAGACGAGCG	8328
Db	8544	GTGTGTGCGACGACTTATGTCGTTATCTGTAAAGTGCGGGGGTTCAGAGAGACGCGCG	8603
Oy	8329	AACTGAAAGCTTTCACGAGGCTATGACAGATTTTGCCCTCTGTGTGACCCCCC	8388
Db	8604	AGCTGAAAGCTTTCACGAGGCTATGACAGATTTTGCCCTCTGTGTGACCCCCC	8663
Oy	8389	AGACCGAGTATGATCTGAGAGCTGATTAATCTTGCTCTCAAAATGTTCTGTGGCGTG	8448
Db	8664	CAACCGAATAGACTTGGAGTTATATACATCATGCTCTCCAAAGTGTCAGTGGCCAC	8723
Oy	8449	GGCCCAAGGCGCCCGCAGATCTACTCTGACAGAGACCTTCACTCCAAATGCCC	8508
Db	8724	GACGCGCTGAAAGAGGGTCTACTACTTACCCGTGACCTTAAACCCCTCGCGAGA	8783
Oy	8509	GCTGCTGGGAAACAGTTAGACACTCCCTGTCAATTTATGGCTGGGAAACATCATCCAG	8568
Db	8784	GCGGCTGGGAGACAGCAAGACACACTCCATGCTTCTGGCTAGGCAATATATCATG	8843
Oy	8569	TACGCCCCGACCATATGGGCTTGCATGCTCTGATGACACACTTTCTCATTTCTCATG	8628
Db	8844	TTTGGCCCCCACTGTGGGCGAGATGATGATGACCCATTTCTTTAGGCTCCTCATTA	8903
Oy	8629	GCTCAAGACAGCTGGACCAAGAACTCAACTTTGATGATGACGAGCGGTATACCTCGTG	8688
Db	8904	GCGAGGATCAAGCTTGAAGACAGGCTTTTAATCTGTGATCTAGAGGCTGCTCACTCATTA	8963
Oy	8689	AGTCCCTTGAACCTCCAGCTTAATATTTAAAGTTACATGGGCTTGAACGCTTTTCTCTG	8748
Db	8964	GAACCACTGGAATCTACCTCCCAATCATTTCAAAGACTCCATGGGCTCAGGCAATTTCACTC	9023
Oy	8749	CACACATACACTCCCAACGAACTGACAGGCTGCTTACAGCCTTCAGAAAATTGGGGCG	8808
Db	9024	CACACTTACTCTCCAGGTGAATTCATATGGGGGCGGATGCTTCAGAAAATTGGGGCTC	9083
Oy	8809	CCACCCCTCAGAGCGTGGAAAGCGCGGACAGTGCAGTCAGAGCGTCTCCTATCTCCGCT	8868
Db	9084	CCGCTCTTCCAGACTTTGGAGAACCGGGGCCGAGACGTCCGCGTCTGTGTCAGA	9143
Oy	8869	GGGGGAGAGCGGCGCTTGGCGGTGATATCTTCAATATGGCGGTGAAGACCAAGCTC	8928
Db	9144	GGAGGCAAGGCTGCTATATATGTGGCAAGTACCTTTCACTGGGCGAGTAAAGAACCAAGCTC	9203
Oy	8929	AAACTCATCTCAATTCGCGAAGCGGCGCTCCGGAATTAATCAGTGTGTTACACGCTGGG	8988
Db	9204	AAACTCATCTCAATATGCGGCGGCTGCGCTGAGACTTGTCCGATGTGTTACACGCTGCG	9263
Oy	8989	GCCGCGGGGCGAGCATTTATCAGACGCTGTGCGCTGCGCAGCCCGCTTATTTGCTCTTT	9048
Db	9264	TACAGCGGGGAGACATTTATCAGACGCTGTCTCATATGCCCGCGCTGCTGTGCTTT	9322
Oy	9049	GGCTCATCTCACTTTTGTAGGGTATGGCTTTTCTTCACTCCCGCTCG	9098
Db	9324	TGCTCATCTCTGCTGCTGCAAGGGGTAGGCATCTACCTCTCCCAACCG	9373

```

RESULT 7
US-08-811-566-5
/ Sequence 5, Application US/08811566
/ Patent No. 6127116
/ GENERAL INFORMATION:
/ APPLICANT: Rice, Charles et al.
/ TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
/ NUMBER OF INVENTIONS: VIRUS (HCV) AND USES THEREOF
/ NUMBER OF SEQUENCES: 21
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: David A. Jackson, Esq.
/ STREET: 411 Hackensack Ave, Continental Plaza, 4th
/ STREET: Floor
/ CITY: Hackensack
/ STATE: New Jersey
/ COUNTRY: USA
/ ZIP: 07601
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/811,566
/ FILING DATE: 03-MAR-1997
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Jackson Esq., David A.
/ REGISTRATION NUMBER: 26,742
/ REFERENCE/DOCKET NUMBER: 1113-1-006
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 201-487-5800
/ TELEFAX: 201-343-1684
/ INFORMATION FOR SEQ ID NO: 5:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 12980 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ HYPOTHEICAL: NO
/ ANTI-SENSE: NO
/
US-08-811-566-5

Query Match 45.4%; Score 4129.6; DB 3; Length 12980;
Best Local Similarity 66.4%; Pred. No. 0;
Matches 6051; Conservative 0; Mismatches 2969; Indels 90; Gaps 5;

QY 1 ATGAGCAAAATCTCTAAACCTTCAAGAAAAACCAAAAGAAACACCAACCGTCGCCCAAA 60
DB 342 ATGAGCAAGAAATCTTAAACCTTCAAGAAAAACCAAAAGAAACCAACCAACCGTCGCCACAG 401
QY 61 GACGTTAGTTTCGCGCGCGCGCGCGCGAGATCGTTGCGGAGATATCTTGTTCGCGCAGG 120
DB 402 GACGTCAGATTCGCGGCGCGCGCGCGCGATCGATCGTTGTGAGTTACTTGTTCGCGCAGG 461
QY 121 GAGCCGAGTTGGGTGTGCGCGCGCAAGAAAGACTTGGAGCGGTCCAGCAGCATGTGA 180
DB 462 GAGCCCTTAATTTGGGTGTGCGCGCGCGCAAGAAAGACTTCCGAGCGGTCCCAACTTGAAGT 521
QY 181 AGGCGCGAGCCCATCTCTTAAAGATGGGCGCTCCACTGGCAATCTCTGGGGAAGAAACCAAGA 240
DB 522 AGACCTCAGCCTATTCCTCCAAAGCAGTGGCCCGAGGCGAAGAACTTGGGCTCAGCCCGGG 581
QY 241 TACCCCTGAGCCCTATATACGGAATGAGGAATCGGCTGGGAGAGATGGCTCTCTGTCCTCC 300
DB 582 TACCTTGAGCCCTCTATATGGAATGAGGAGTGGCGGATGGGCGGAGATGACTCTGTCTCC 641
QY 301 CGAGGTTCCGCTCTCTTGGGAGCCCAATGATCCCCCGCATAGTTCGCGCAACGTGGGT 360
DB 642 CGTGGCTCTGAGCTGAGGCGCCCAAGACAGCCCCGCGCGTGAAGTTCGCGCAATTGGGT 701

```


Qy	361	AAAGTCATCCGATACCCTTAACGTCGGGCTTTGCCACCTCAAGGGGTAAATCCCTGTGCTG	420
Db	702	AAAGTCATCCGATACCCTTAACGTCGGGCTTTGCCACCTCAAGGGGTAAATCCCTGTGCTG	761
Qy	421	GGCCGCCCGCTCGGCGGGCGTGCAGACCTCTCGGCAATGGCGATAGAGTCTTGAGAGAC	480
Db	762	GGCCGCCCGCTCTTGAGAGGGCTGCGAGGGCCCTGGGGCATGGCGTCCGGGTTCTGGAAAC	821
Qy	481	GGGGTTAATTTTGCACAGGGAATTACCCGGTTGCTCTTTCTATCTTTCTTGCGCC	540
Db	822	GGCGTTAACTATGCAACAGGAACCTTCTGGTGGCTCTTCTCTATCTTCTTGCGCC	881
Qy	541	CTGCTGATCTGCAATACCAACCCCGGCTCCGCTGCGGAAGGAATCATGATACGGC	600
Db	882	CTGCTCTCTTCCTGACCGGTGCGGCTTCAAGCTACCAAGTGGCAATCTTCGAGGCTT	941
Qy	601	TACATGGTGACTAACGACTGCAACCATAGACATTAATCTGGCAGCTTCAGGCTGCTG	660
Db	942	TACATGTCACCAATGATTTGGCTTAATCGAGTATTTGTAACGAGCGCGGATGTCATC	1001
Qy	661	CTCACAACCTCCCGGCTGCTCCCGTGCAGAAAGTGGGGAATGCATCTCAATGCTGATTA	720
Db	1002	CTGCAACATCCGGGGTGTGTCCCTTGGCTTTCGAGGGGTAAACGCTCTCAGGTGTTGGTGG	1061
Qy	721	CCGGTCTCAGCGAAATGTGGCGGTGTCAGGGGCGCGGCGCTTCACGACAGGGCTTGGAGG	780
Db	1062	GGGGTGAACCCCAACGGTGGCCACAGGGAACGGCAAACTCCCAACACGCACTTGCAGT	1121
Qy	781	CACATTCGACATGTTGTGATGTCCGCCACGCTCTGCTCTGCCCTTCAATGTGGGGACCTC	840
Db	1122	CATATTCGATCTGCTGTGTGCGGAGGGCCACCTCTGCTCGGGCCCTTCAATGTGGGGACCTG	1181
Qy	841	TGGCGTGGGGTGATGCTGCGAGCCCAATGTTCAATGTCCTGCGCGACACACAATGCTTT	900
Db	1182	TGGCGGTCGTCTTTCTTTGTTGTGTCACATGTTTACCTTCTCTCCAGGCGGCACTGACG	1241
Qy	901	GTCGAAGACTGCAAATGCTCCATCTAACCTGTATACATCACTGGAACAACGATGGACATG	960
Db	1242	ACGCAAGACTGCAAATGTTCTATCTATCCGGCCATATACGGGTCATTCGATGGCATGG	1301
Qy	961	GACATGATGATGAACCTGGTGGCCCAACGGCTACATGATCTTGGCGTACGCATGTGTCT	1020
Db	1302	GATATGATGATGAACCTGGTCCCTCAACGGCACGTTGGTGTAGTCACTCACTCTCCGATTC	1361
Qy	1021	CCCGAGGTCATTAATAGACATCATTAAGCGGGGCTCATTAAGGGCGTCAATTTGGCTTGACC	1080
Db	1362	CCACAAGCATCATGACATATGGCTGTGTCTCACTGGGGAATCTTGGCGGCAATAGCG	1421
Qy	1081	TACTTCTCTATGACAGGAGCGTGGCGGAAATGCTGTTCATCTTTCTGTTGGCCGCGGG	1140
Db	1422	TATTTCTTCATGATGGGGGAACTGGGCGGAAGTCTCGTAGTGTCTGTATTTTGGCCGGC	1481
Qy	1141	GTCGACGGGCGACCCCAATACGTTGGGGGGTTCTGCGCGGACAGACACCGGGCGGCTACAC	1200
Db	1482	GTCGACGGGCGAAACCCACAGTCAACGGGGGAAATGTGCGGCGGACCAACGGCTGGCTGTT	1541
Qy	1201	AGCTTATTTGACATGGGGCCCGAGCAGAAAAATCCAGCTCGTTAAACCAATAGCAGCTGG	1260
Db	1542	GGTCTCTCTTACACAGGGGCGCAAGAGAAATCCAACTGATCAACACCAACGGCAGTTGG	1601
Qy	1261	CACATCAACCGCACCGGCTTGAATCTGCAATGAATCTCTTGCAACCGGCTTATAGCGCTCT	1320
Db	1602	CACATCAATACACGGCTTGAATCTGCAATGAATGAAGCTTAAACACCGGCTGTATGACAGG	1661
Qy	1321	CTGTGTTCAACCAACGCTTCAACTCGTCAAGATGTCCGAAACGATATCCGCTCGCGC	1380
Db	1662	CTCTTCTATCAACCAATTAACCTTCAAGGCTGTCTCGAAGGTTGGCAGGCTGCCGA	1721
Qy	1381	AGTATCGAGGCGCTTCCGGGTGGAGTGGGGCGCTTGCAATATGAGAGTAAATGTCAACCAAT	1440
Db	1722	CGCCTTACCAATTTTGGCCAGAGGCTGGGGTCTTATCAAGTATATGCCAAAGGAAGCGGCTCTC	1781

QY	1441	CCAGAAGATATGAAACCTTATGCTGGCACTACCAACGAAGGAGTGGCGTGCCTC	1500
Db	1782	GACGA-----ACGCCCTTACTGTGGCACTTACCCTTCMAAGACTTGTGGCAATTGTGCC	1835
QY	1501	GCGAAGACTGTGTGTGGCCGAGTGTACTGTTTACACCCCGACGCCAAGTGTAGTGGACG	1560
Db	1836	GCAAAAGACGTGTGTGGCCCGGTATATTTGCTTCACTCCGACCCCGTGTGTGGGAACG	1895
QY	1561	ACCGACAGCTTGGAGGGCCACTTAAACGTGGGGGAGAAATGAGACAGATGTCTTCCTA	1620
Db	1896	ACCGACAGGTGTGGCGCGCTTACTTAAGCTGGGGGTCAAAATGATAGAGATGTCTTC	1955
QY	1621	TTGAACAGCACTGCACCAACCGCTGGGGGTCAATGTTCGGCTGCACGTGTGAATGAATCTTCT	1680
Db	1956	CTTAAACAACACAGGCCACCGCTGGGAATTTGTGTGGTGTGTACTGTGAATGAATCACT	2015
QY	1681	GCGTACACCAAGACTTGGGCGGCACACCCCTGCCGTACTAGAGTGACTTAAACGCCACG	1740
Db	2016	GGATTTCAACAAGGTGCGGAGGCGCCCTTGTGTCTACCGAGGGGTGGGCAACACA--	2073
QY	1741	ACGGAACCGTTGTGGCCCAAGAGTGTTTAAGGAAGATCTGATATCACTTACCTTCAAA	1800
Db	2074	----CTTGTCTGTGCCCACTGATTTGTTCCGACAGCATCCGAAGCCAATACTCTCGG	2129
QY	1801	TGCGGCTCTGGGCCCTGTGGCTCAGCGCCAAAGTGCCTGATTCAGTACCCTTACAGGCTCTG	1860
Db	2130	TGCGGCTCTCGGTCCCTGTGATTTACACCGAGGTGATGTGTGACTACCGTATTAAGCTTTG	2189
QY	1861	CATTACCCCTGCACAGTTAACTATACATCTTCAAAATAAGATATATATGTGGAGGGGTT	1920
Db	2190	CACATATCTTGTATCAATCAATCAATTAACCATATTAACAATGACAGATATGACGTGGAGGGGCT	2249
QY	1921	GAGCACAGGCTCAGCGGTGCATGCATTAATTCATCTGTGGGAGATGTTGCAACTTGGAGAC	1980
Db	2250	GAGCACAGGCTGGAAGGGCGCTGGAATGGAACCGGGGCGAAGCTGTGATCTGGAAAGAC	2309
QY	1981	AGAGACAGAACTCACTGTCTCTTTGTGTGACTCCACACGGAATGGGCCATTTTAACT	2040
Db	2310	AGGACAGAGTCCGAGCTCAGCCCAATGTCTGTCTGTCCACACAGTGGCAGGTCTTCCG	2369
QY	2041	TGCTCTTAACGTGAGCCGCGCGCTGTGTGACGTGTCTTCACACCTCCACAAACATC	2100
Db	2370	TGTTCTTTCACAGACCTTCGACGAGCTGTTCACACGGCTCATCACTCCACACGAACATT	2429
QY	2101	GTGACGTATCAATTCATGTATGTGCTATCACTGTCCCTCCAAATAATCATCTGTCCGATG	2160
Db	2430	GTGACGTATGACATCTTGTATCGGGGTATGGGTCAAGATGTGCTGTGGGCTAATAAGTGG	2489
QY	2161	GAGTGGGTAACTCTTATTTCTGTCTTTAAGCGGAGGCCAGGGTTTGTGGCTGTATATGG	2220
Db	2490	GAGTACGTGTCTTCCGTGTCTCTGTCTTCGACAGACGCGGCGTGTGTCTCTCTGTGTGG	2549
QY	2221	ATGTCTATCTTGTGTGGGCCAAGGCGGAAGAGCACTAGAGAAGGTGTGCATCTTGGACGT	2280
Db	2550	ATGATGTATCTCATATATCCCAAGCGGAGGCGGCTTTGGAGAAGCTCGTAATATCTAATGCA	2609
QY	2281	GCGAGCGAGCTAGCTGCAATGTCTTCTATATTTTGTATCTTTTTTGTGTGGTCTGTTGG	2340
Db	2610	GCATTCCTCGCGCGGGAAGCAACGCTGTGTGTCTTCTCTCGTGTCTTCTGTGTGGCTGG	2669
QY	2341	TACATCAAGGGTGGGGTAAATCCCTTATAGTAACTATTTCCCTCACTGGCCTGTGTCTT	2400
Db	2670	TATCTGAAGGGTAAAGTGTGTGTCCCGAGCGGTCTTACGCTTTTACAGGGAATGTGGCTCTC	2729
QY	2401	AGCTTACTGTCTCTAGCAATGGCCCAACAGGCTTATGTATAGACGATCTGTGTGATGC	2460
Db	2730	CTCTGTCTCTGTGGGTGTGCTTCAACGCGGCATTAAGCACTGAGACAGAGGTGGCGCG	2789
QY	2461	CAGATTAAGAGCGGTCTGTGTGTATATATATCTCTTTATCTGTCAACCCCGGGTATTAAG	2520
Db	2790	TGTGTGTGCGCGGTTGTCTTGTGTGTGGGTATATAGCGCTGATCTGTGTGCCATATTAACAG	2849
QY	2521	ACCTTCTCAACCGGTTTTTGTGTGTGTGTGTGCTATCTTTGACCCCTGTGGGGAAGCTATG	2580

Db 2850 CGCTACATCAAGCTGTGATGTGGCTTCAAGTATTTCTGACAGAGTAAAGCCCAA 2909
QY 2881 GTCCAGAGTGGGACCAACCTATGACAGGTGGCGGTGGCGTGTATGACATCATATGAGGC 2640
Db 2910 CTGACAGTGTGGATTTCCCTCCCAAGCTCCGAGGAGGAGCGATGCGCATTTTATCTC 2969
QY 2641 GTCCCATATTTCTACCCAGGTGTGTGTTGACATTAACCAAGTGGCTTTGGCGGTGCTT 2700
Db 2970 ATGTGTGTGTACACCGACTCTGGTATTTGACATACCAAACTACTCTGGCCATCTTC 3029
QY 2701 GGGCTGTGTAACCTCTTAAAGGTGCTTTGACGGCGGTGCTTCTTCTGAGGGCTCAC 2760
Db 3030 GGACCCCTTTGAAATTTTCAAGCCAGTTTGTAAAGTCCCTTACTTCTGCGCGTTCAA 3089
QY 2761 GCTCTACTGAGATGTGACCACTGACCAAGCATCTCGGGGGGACAGTACGTCCAGATG 2820
Db 3090 GGCCTTCTCCGGATCTGGCGCCTAGCCGGAAAGTACCGGAGGTCACTATAGTGCAAATG 3149
QY 2821 GCGCTACTAGCCCTTGGAGGTGACCTGACCTTACATTAATGACCACTTCACTTATG 2880
Db 3150 GCCATCATCAAGTTAGGGGCGCTTACTGGCACCTATGTATTAACCATCTCACTCCCTCTT 3209
QY 2881 TCGGATTTGGGTGCTAGTGGCTGCGGAGCACTGGCGGTGCGCCTTGAAGCCTATATCTTC 2940
Db 3210 CGAAGCTGGGCGCAACAGGCTCGAGATCTGGCGGTGTGGAACCAAGTGTCTTC 3269
QY 2941 AGTCGATGAGAGAAAGTCAATGTCTGGGAGCGGAGACAGTGTGTTGGGGACATT 3000
Db 3270 TCCGAATGAGACCAAGCTCATACGTGGGGGACATACCGCGCGGTGGGTGAATC 3329
QY 3001 TTACACGAGACTTCCGTGTCCGCGGACTTGTGTGGAGGTCTCTTGGCCAGCTGAT 3060
Db 3330 ATCAACGGCTTGCCCGTCTGCGGTAGGGGACAGAGATATGCTTTGGGCCAGCGAC 3389
QY 3061 GGCTATACCTCCAGAGGGGTGAGTCTTCCGCCCATCACTGTTTACGCCACAGACA 3120
Db 3390 GGAATGTCTCCAAAGGGTGAAGTGTGGCGCCATCAAGGCGTACCGCCACCAAGC 3449
QY 3121 CGTGGCCCTTTGGGACCATAGTGGTGAAGTGAACGGGGCGGACAAAGACAGACAGCT 3180
Db 3450 AGAGGCCCTTCAGGGGTGATATACACAGCTGACTGGCCGGGACAAAACCAAGTGAAG 3509
QY 3181 GGGGAAATTCAGTCTGTGCAAGTCACTCAAGTCTTCTCGGAACTCCATCTCGAGG 3240
Db 3510 GGTGAGGTCCAGATCTGTGTCAACCTGTACCCAACTTCTCGGCAACGTGATCAATGGG 3569
QY 3241 GTTTTGTGAGCTGTCTACCATGAGTGGCAAAAGTCTGGCCGGGTCAAGGGGTGCG 3300
Db 3570 GTATGCTGAGCTGTCTACCGAGGCGGGAACGAGGACATCGATCACCCAAAGGTCTT 3629
QY 3301 GTCAAGCAGATGTAATCTCAGTGTGAGGGGAGCTTATGAGGGGTGAGCCCGCCCTGGG 3360
Db 3630 GTCAATCCAGATGTATACCAATGTGAGCAAAAGCTTTGTGGGTGGCCCGCTCTCAAGT 3689
QY 3361 ACTAATCTTTGAGACCGGTGACAGTGTGAGCGGTGCACTGTACTGTGTGTCACGCGAAGC 3420
Db 3690 TCCGCTCATTTGACACCTGTGACCTGTGCGGCTCTTGGACCTTTACTGTGTCAAGAGGAC 3749
QY 3421 GCTGATGTATCTCCGGCTCGAAGAGCGGGGACAAACGGGAGAGGCTTACTTCCCGAGA 3480
Db 3750 GCGGATGTCAATTCGCGCGCGCGGAGAGTATACAGGGGTGAGCTGTGCTTCCCGCGG 3809
QY 3481 CCTCTTTCACCTTGAAGGGGTCTCAGAGAGCCCGGTGTATGCCCCAGGGGACAGCT 3540
Db 3810 CCGATTTCTACTTAAAGGCTCTCGGGGGGTCCGCTGTGTGTGCCCCCGGGACAGCC 3869
QY 3541 GTCCGAGTCTTCCGGGACAGTGTGTCTCTCGGGCGGTGAGTCAATGATATTTATC 3600
Db 3870 GTGGGCTATTTCAAGGCGCGGTGTGACCGGTGAGTGGCTTAAGGCGGTGACATTTATC 3929
QY 3601 CCGGTGAGACACTGACATGTGACGCGGTGCCCACTTATGAGCAACAGCACCA 3660

Db 3930 CCTGTGAGAACCTAGAGCAACCATAGATCCCGGTGTTTACCGGACAACTCTCTCCA 3989
QY 3661 CCTGTGTGCCAGACCTATGAGTGGGTACTTGTGATGCCCGGACTGGCAGTGAAG 3720
Db 3990 CAGACAGTCCCGGAGCTTCCAGGTGGCCGACGTGACGTCTCCACCGGACGGGTAAAG 4049
QY 3721 AGCACAAAGTCTGTGCGATATGTGCTCAGGGGTATTAAGTGTATGTCTTAATCCC 3780
Db 4050 AGCACCAAGTCCCGCTCGTACGAGCCGAGGCTTACAGAGTGTGTGTCTCAACCCC 4109
QY 3781 TCAGTGGTCCGACCCCTGGGGTTTGGGGGCTCTTGTCTAAGGACATGATCAATCCC 3840
Db 4110 TCTGTGTGCAACGCTGGGCTTTGGCTTTACATGTCCAAAGGCCATGGGTGATCTCT 4169
QY 3841 AACATTAGACTGAGTCAAGACTGTGACGACCGGGGCGCCATCACTGACTCCATAT 3900
Db 4170 AATATCAGAACCGGGGTGAGAACATTAACATCGGAGGCCCATCGTACTCACCTAC 4229
QY 3901 GGCAAATTCCTGGCGATGGGGCTGTGGCGGCGGCTTACGACATCATATGTGTAT 3960
Db 4230 GGCAGTTCCTTGCAGCGCGGGTGTGAGAGGTCTTAAGACATTAATTTGTGAC 4289
QY 3961 GAATGCCATGCGGTGACTTACCAACATCCTTGGCATGGAACAGTCTTGATCAAGA 4020
Db 4290 GAGTGCACCTCAAGATGCAATCATCTTGGCATGAGCATGTCTTGAACAAAGCA 4349
QY 4021 GAGACAGTGGGTGAGCTTAATCTGTGTGCTACAGCTTACGCCCCCTGGGTCAGTGACA 4080
Db 4350 GAGACTGGCGGGCGAGACTGGTGTGTGCGCACTGTACACCTCCTCGGACTCCGTCAT 4409
QY 4081 ACCCCCAACCCAACTAAGAGAGTGGCCCTTGGCAGAGAGGCGAGATCCCTTCTAT 4140
Db 4410 GTGTCCATCTTAACATCAAGAGGTGTGCTGTCCACCAAGGAGATGCCCTTTTAC 4469
QY 4141 GGGAGGGGCAATCCCTGTCTTAATCAATCAAGGAGAGAACATGTGATCTTGGCATTA 4200
Db 4470 GGCAGAGCTATCCCTCTCAAGGTATCAAGGGGAGAAACATCTACTTCTTGCACCTCA 4529
QY 4201 AAGAAAGTGTGACAGAGTCTCGCGGCGGCTTCCGGGATATGGGCTTGAATCAAGTGA 4260
Db 4530 AAGAAAGTGTGACAGAGTCTCGCGGAGGTGTGTGATTTGGGCTCAATGCGGTGGCC 4589
QY 4261 TACTACAGAGGTGAGAGTCTTCCGTAATTAACCACTCAGGAGAGCTATGTGTCTGCGC 4320
Db 4590 TACTACCGGGTCTTACGATGTCTGTATCTCCGACCAAGCGGAGTGTGTGTGCGTGTG 4649
QY 4321 ACCGAGCCCTCATGACAGGATTAATCTGGGGAATTGATCCGATGTCAGTCAAGCTGA 4380
Db 4650 ACCGATGCTCATGACTGTGCTTTACGGGACTTGACTGTGTATGACTGCAACAG 4709
QY 4381 GCGGTCACTCAAGTTGATGACTTCAAGTTTGAACCCCAATTCACATTAACCAACAGATT 4440
Db 4710 TGTGTCACTCAGACATGTGATTTCAAGCTTGAACCTTAACATTAAGAACACAG 4769
QY 4441 GTCCCTCAGAGCCTGTCTGACATGACAGCGCGGGGTGTGCGAGGGTATGGGAAAGACTG 4500
Db 4770 CTCGCCAGAGTGTGTCTCTCAGAGCTCAACCGCGGGGACAGACTGCGAGGGGAAAGCA 4829
QY 4501 GGCATTATATGATGATTTTCACTGTTGAGCGAGCTCAGGAATGTTTGAACATGTATGTG 4560
Db 4830 GGCATCTCAAGATTTGTGTGACCTGGGAGAGCGCCCTTCGCGCATGTTCATGTCTCGCTC 4889
QY 4561 CTCTGTGAGTGTACAGACGAGGGGCGGATGTATGAGTCAACATCGAGAGACACC 4620
Db 4890 CTCTGTGAGTGTATGACGCGGCTGTGTGTATGAGTCAACGCGCGGACGAGACTAC 4949
QY 4621 GTCAAGTCTAGGCGGTATTTCAACAGCGCGGCTTGTGCTGTGCGCAAGACATCTTTGAG 4680
Db 4950 GTTAGGCTACAGACGTATGAACACCGCGGGCTTCCGTGTGCGCAGGACCATCTTTGA 5009
QY 4681 TTTTGGAGGCAATTTTCAACCGGCTCAACACATATGATGCGCACTTCTTCCAAAC 4740
Db 5010 TTTTGGAGGCGCTTTTACGGGCTCACTCATATGATGCCCACTTTTCTATCCAGACA 5069

4741 AAGCAATCGGGGAAATTTGCACTACTTAACAGCCCTACAGGCTAGAGTGGCTAGG 4800
5070 AAGAGAGTGGGGAACCTTCTTACCTGGTAGCAAGCACCGTGTGGCTAGG 5129
4801 GCCAAGCCCCCCCCGCTCTGGAGCTCATGTGAAGTGTGACTGCACTCAAGCCC 4860
5130 GCTCAAGCCCCCTCCCATCGTGGAGCCAGATGTGGAAAGTGTGATCCCGCTTAAACC 5189
4861 ACACTCGTGGGCCCCCAACCTCTCTGTGACCGCTTGGGCTCTGTTCACAGAGTCAAC 4920
5190 ACCCTTCATGGGCGCAACCCCTGCTATACAGACTGGGCGCTGTTCAGAAATGAAGTCAAC 5249
4921 CTCACATCCCGGAGGAAATATACGCGCACCTGACAGAAAGGCGACTTGGAGTATG 4980
5250 CTGACGACCCCAATACCAATATACATATGACATGATGCGCGACCTGGAGAGTGTG 5309
4981 ACCAGCATGAGGCTTGGGAGGGGAGTCTTGGCGGCGCTCGCGGCTATTTGGCTGCG 5040
5110 ACAGGACCTGGGCTGCTGGTGGGGGCTCTGGCTGCTGCGCTGATATGCTGTCA 5369
5041 ACCGGTGTGTGTCATCATGCGCGCTTGCACATTAACAGCGAGCCGTGTGGCGCG 5100
5370 ACAGGCTGCGTGTGATAGTGGGAGGATTTGCTTGTCCGGGAGCCGGGCAATTATACT 5429
5101 GACAGGAGGTCCTCTATAGAGGCTTTGATGAGATGAGGAAATGGCTCTAGGGCGGCT 5160
5430 GACAGGAGGTTCTTACAGAGAGTGTGATGAGAGAGAGTGTCTACAGCACTTAACG 5489
5161 CTCATTGAGAGGGGAGCGAGTAGCCAGATGCTGAGTCCAGATCCAGAGCTTATG 5220
5490 TACATCGAGAGGAGATGATGCTGCTGAGAGTTCAAGCAGAGGCGCTCGGCTCTG 5549
5221 CAGCAGCTTCCAAACAGCTCAAGACATACACCACTGTGAGGCTTCATGGCCCAAG 5280
5550 CAGACCCGTCGCCCGCAAGGAGATATCACCCCTGCTCCAGACCAACTGGCAAAA 5609
5281 GTAGAACATTTGCGGCGCAACACATGTGGAATTCATAGCGGCACTCAATACTGCA 5340
5610 CTGAGGTCTTGTGGGAGAGACATGTGAAATTTCAAGTGGAGATACATATCTGCG 5669
5341 GCACTATCAACACTGCGAGGGAACCTGAGTAGCTTCCATGATGGGCTCAGTGCCGC 5400
5670 GCGCTGTCAAGCTGCTGTGTAACCCCGCATTTGCTTATGATGGCTTTTACAGCTGCC 5729
5401 CTACACAGTCCGCTGTCAACAGCAACATATCTTCAACATTTTGGGGGCTGTGCTA 5460
5730 GTACACAGCCCACTAACCACTGGGCCAAACCTCTCTTCAACATATTTGGGGGCTGGGTG 5789
5461 GCATCCCAATTTGACACACCGCGGGGCACTGGCTTTCGTTGTCAGTGGCTATAGTGA 5520
5790 GCTGCCAGCTCGCGGCGCCCGGCTGCGCTACCGCTTTGTGGCGCTGAGCTGAGGC 5849
5521 GCTGCCGTAGGAGATAGGCTTAGTAAAGTGTCTAGTGAACATCTGGAGAGGATAGT 5580
5850 GCGGCCATCGGAGAGCTTGGAGCTGGGAGAGTCTCTGTCGACATTTTGGAGGATAGGC 5909
5581 GCGGCAATTTTGGGGGCTCTGTGTGCAATTAAGATCATGTCTGGCGAAGCCCTCATG 5640
5910 GCGGCGCTGAGCGGAGCTCTTGTAGCTTCAAGATCATGAGCGGTGAGGCTCCCTCCACG 5969
5641 GAGGATGTGTCACCTTGTGCGCTGGAATTTGTCTCGGGTGTCTTGGTATGAGTGGAGTC 5700
5970 GAGGACTGTGTCATTTGTGTGCGCGCATCTCTGCGCTGAGGCTTTGTATGTGGTGTG 6029
5701 ATCTGCGGCGCAATTTGTGCGCGCAACAGTGGGACCGGGGAAAGCGCGCTTCAATGATG 5760
6030 GTCTGCGGCAAGATCTGCGCGGCAAGTGGGCGCGAGGCGAGAGTGCATGATG 6089
5761 AATGACTCATTTGCTTTGCTTCCAGAGGAATCACTGCGCCCACTACCTAGTACG 5820
6090 AACCGGCTAATAGCTTGTGCTCTCCGGGGGAAACATATTTCCCCACGCACTAGTGTGCG 6149

5821 GAGTCGAGTCTGCGAGCGGTGTGACCACTACTTGGCTCCCTTACATTAACAGGCTG 5880
6150 GAGAGGATGACAGCCCGCGCTCATCTGCACTACTACAGAGCTTCACTGTAAACCACTC 6209
5881 CTGAGAACTCCCAACTGATTAATGAGAGTCCCGCATTCATGCGGCGCTGTGG 5940
6210 CTGAGGCGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6269
5941 CTGCGGATGTGTGGAGCTGGGTTTGACCATCTTAACAGACTTTTAAATTTGGCTGACC 6000
6270 CTAAGGACATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6329
6001 TCCAAATTTATCCCAAGATGCGCGGCTCCCTTTGTCTCTGTCAAAAAGGGTACAG 6060
6330 GCGAGCTCATGCAACACTGCTGGATTCCTTTGTGTCTGTGCGAGCGCGGATAGG 6389
6061 GCGGTGTGGCGGCACTGAGCATGACCAACGCTGTCTTGGCGGCAATATCTCT 6120
6390 GGGGTCTGCGAGAGAGACGCGCATTTATGACACTGCTGCGCACTGTGGAGCTGAGATCACT 6449
6121 GCGAATGTCGCTTGGGCTCCATGAGAAATGACGGGGCTTAAAGACCTGCAATATCTG 6180
6450 GGCATGTCAAAAACGAGAGATGAGAGATGCTGTGTCTAGACCTGACAGAAATGTG 6509
6181 CAGGGACCTTCTCTATCAATTTGTACACGAGAGGCGAGTGTGCGGAAACCGCGCA 6240
6510 AGTGGAGCTTCCCATTTAAGCTTACACAGCGGCGCTGTACTCCCTTCTGTGCGCG 6569
6241 AACTTTAAGTGTGCACTGTGAGAGTGTGCGGCTCAGAGTACGCGAGGTGACGACGAC 6300
6570 AACTATTAAGTGTGCGGCTGTGAGAGGTGTGCGAGAGAAATATGAGAGATTAAGCGGAGT 6629
6301 GGGTCATACCATCACTAACAGACTCACCACTGATTAATCTTGAAGTCCCTCCCACTA 6360
6630 GGGGACTTCCACATGATGTGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 6689
6361 CCCCTTCCGAGTCTTCTTCTGAGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 6420
6690 CCAATGCGGAAATTTTTCACAGAAATGTGAGCGGGTGTGCTACATAGTGTGTGCGCCCT 6749
6421 CCGAAGCGTGTCTTCTGCGGAGTGTGCTGTGCTGTGCGGCTTAAATTTGTGTGCT 6480
6750 TGCAGGCTTGTGCTGCGGAGGAGTATCATTTAGAGTATGAGTATGAGTATGAGTATGAGT 6809
6481 GGGTCCAGCTTCTTGTGCAACCTGAAACCGGACGACAGAGTATGATGATGATGATGATGATG 6540
6810 GGGTCCCAATTTACTTGTGCGAGCCGGAACGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 6669
6541 GATCCATCTCATATCAACGCGGAGAGCTGAGCGCGGCTTAAAGCGGCGGATGACCCCA 6600
6870 GATCCCTCCATATTAACAGAGAGGCGCGCGGAGAAAGTGTGCGAGAGAGGATCAACCCCT 6929
6601 TCCGAGCAAGCTCTTGTGCGAGAGCTATGCGCAACATGTGCGAGGCACTGTGAC 6660
6930 TCTATGCGAGCTCTTGTGCGAGAGCTGTGCGCTCATCTCTCAAGGCAACTTGTGAC 6989
6661 ACCCAGCGCAACCTATGATGTGAGACATGTGTGATGATGATGATGATGATGATGATGATG 6708
6990 GCCAACCATATCTCCCTGAGCGCGAGCTCATAGAGCTTAACTCTGTGAGAGCGAGAG 7049
6709 ATGGGGGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6768
7050 ATGGGCGGCAACATCAACAGAGTGTGAGTCAAGAAACAAAGTGTGATTTCTGAGTCTTCT 7109
6769 GACCCATGTGCAAGAAAGAGAGCACTTGAAGCTTGTGATACATCAATCAATCAATCAAT 6828
7110 GATCCCTGTGTGCAAGAGAGATGAGCGGAGATGATGATGATGATGATGATGATGATGATGATG 7169
6829 CCCAAGAGAGTGTCCCAACAGCTTAAAGCGGCTGTGAGACAGGCTGATTTTCAACCAACG 6888
7170 AAGTCTGAGATTTGCGCGGCGCTGCGGCTTGTGGCGGAGCTTCAACACCCCGCG 7229
6889 CTGTGGAATCTGTGAAAGAGCGAGATTAACCAACCGGCACTGTGTGCGGCTGTGCTCTC 6948

Db 7230 CTAGTAGAGAGTGGAAAAAGCTGACACCAACCTGTGTCTCAATGGCTGCCCCCTA 7289
Oy 6949 CCTCTCTTAGAAAAACCCCGAGCGCTCCCAAGAGAGCGCGGACAGTGGGCTAAGT 7008
Db 7290 CCACCTCCACGAGTCCCTCTCTGTGTCTCCGCTTGAAAAAGCTTACGAGTGTCTAC 7349
Oy 7009 GAGGACTTCATAGAGATGCCCCCTTCAACAGCTGGCCATTAACTCTTTGGCAGCCCCC 7068
Db 7250 GAATCAACCTATCTACTGCTTGCGGAGCTTGCCACCAAAAGTTTGGAGCTCTCTCA 7409
Oy 7069 CCAAGCGCGGATTCAAGGCTTTTCCACGGGGCGGGCGCTGCCATTCCGGCAGTCAACG 7128
Db 7410 ACTTCGGGCATTACGGGGGCAATACGACATCTCTTGAGCCCGC----- 7457
Oy 7129 CCTCTGATGATGTGGCCCTTTTCGAGACAGGTTTCAATCTTTCCATGCCCCCTCGAG 7188
Db 7458 CTTCTGGCTGCCCCCGGACTCCGAGCTTGAAGTCTATTCTTCAATGCCCCCTTGAG 7517
Oy 7189 GGGGAGCTTGAGATCCAGACTTGAGCCTGAGAGGTAGAGCCCAACCCCGCCGAG 7248
Db 7518 GGGGAGCTTGAGATCCGGA----- 7537
Oy 7249 GGGGGGGTGGAGCTCCCGGCTCGGACTCGGGGTCTGTGTCTACTTGTCCGAGAGAC 7308
Db 7538 -----TCTCAGGAGCGGATCATGTGTGACGCTCAAGTGTGGGCGGACAG 7583
Oy 7309 GACTCCGTGTGTGTGCTGCTCATGCTCATCTCTGAGACGGGGGCTTAATACTCTTGT 7368
Db 7584 GAAAGTGTGTGTGTGTCTCATATGTCTTAATCTCTGAGAGGCGACATCTGACCCCTGC 7463
Oy 7369 AGTCCCGAAGAGAGAACTTACCGATTAACTCTTGAGCAATCTCTGTGGATATCAC 7428
Db 7644 GCTCGGGAAGAAACAAAACCTGCCATCAACGACATGAGCACTGTTGCTACGCGCATAC 7703
Oy 7429 AACAAGTGTACTGTACCAACAAGAGCGCTCATCTAAAGGCTAAAAAGTAACTTTT 7488
Db 7704 AATCTGTGTATTTCCACCACTTCAAGAGTCTTGCCAAAGCAGAAAGTACATTT 7763
Oy 7489 GATAGGATGCAAGTCTGACTCTTCACTCACTCACTCACTTAAAGCATTTAACTTACG 7548
Db 7764 GACGACTGCAAGTCTTGAAGACCAATTAACGAGAGCTGTCMAAGAGTAAAGCAGCG 7823
Oy 7549 GCCTTCAAGGTCAACGCAAGGCTCTCACTGAGAGAGGCTTCCAGTTAAACCCACC 7608
Db 7824 GCGTCAAAAGTGAAGGCTAACTTGTCTATCCGTAGAGGAAGCTTGACGCTGACCCCA 7883
Oy 7609 CATTCGCAAGATCTAAATATGGGTTTGGGCTAAAGAGTCCGAGCTTGTCCGAGAG 7668
Db 7884 CATTCAGCCAAATCCAAAGTTTGGCTATGGGCAAAAGAGCTGCTGCTGATGCAAGAG 7943
Oy 7669 GCGGTTAACCATGATCTGCTGAGAGAGCTCTGAGAGACTCAGAAACCAATTT 7728
Db 7944 GCGGTAGCCCATCACTCCGTGTGAAAGACCTTCTGGAAGACAGTAAACCAATA 8003
Oy 7729 CCCAACCATTAATGCGCAAAATAGAGTGTCTGCTGGAACCCCAACCAAGGGGGCAAG 7788
Db 8004 GACATTAACATATGSCCAAGAACGAGTTTCTGCTTCAAGCTGAGAAAGGGGGTCTGT 8063
Oy 7789 AAAGAGCTGCGCTTATGCTTAAAGCTTCCGCGCTCAGAGTCTGGAAGAAATGACC 7848
Db 8064 AAGCAGCTGCTGCTATGCTGTCTCCGACCTGGGCGTGGCGGTGGAAAGATGACC 8123
Oy 7849 CTTTATGACATTACAAAACTTCTCAGGCGGTGATGGGGGCTTCTTATGATTTCCAG 7908
Db 8124 CTGTACGACGTGTGTAAGAGTCCCGCTGGCGGTGATGGAGTCTTACAGATTTCCA 8183
Oy 7909 TATTCCTCCGCTCAGCGGCTAGAGTTTCTTTGAAAGCATGCGCGAAAGAGACCTT 7968
Db 8184 TACTACAGGACAGCGGCTTGAATTCCTCTGTGCAAGCGTGGAGTCCAAAGAACCCCG 8243
Oy 7969 ATGGTTTTTGTATGATACCGGATGCTTGAATCAACCTGATCTGAAGAGACATCCAG 8028

Db 8244 ATGGGTTCTGTATGATACCCGCTGTTTGACTCCACAGTCACTGAGAGCATCCGT 8303
Oy 8029 ACTGAGAGTCAATATATGAGGCTGCTCTTGCCCGGAGAGGCCACACTGCCATAC 8088
Db 8304 ACGAGAGAGCAATTAACAAATGTGTGACCTGACCCCAAGCCCGCTGACATCAAG 8363
Oy 8089 TCGTAACTGAGAGCTTTACGTGGAGGGCTTATGTTCAACAGCAAGGCCAAACTTGC 8148
Db 8364 TCCCTCACTGAGAGCTTTATGTTGGGGCCCTTTACCAATTCAAAGGGGGAAAACTGC 8423
Oy 8149 GGGTACAGCGTTGGCCGCGCAGCGGGGCTCACACACTAGATGGGGAAACCAATCACA 8208
Db 8424 GGTACCGAGAGGTGCGCGCGAGCGGCTACTGACAACTAGCTGTGTAACCTCACT 8483
Oy 8209 TGTACTGAAAGCTTTAGCGGCTTGTAAAGCTGACAGGATATACGCCCAATGCTG 8268
Db 8484 TGTATCATCAAGGCCCGGGCAGCTGTGAGCGCGAGGGCTCCAGGACTGACCATGCTC 8543
Oy 8269 GTATGCGGATGATCTTGTGTCTCATGAAAGCCAGGGGACGAGAGAGACGAGCGG 8328
Db 8544 GTGTGTGGGAGCATTAAGTCTGTATCTGTGAAGGTGCGGGGCTCCAGAGAGACGCGGCG 8603
Oy 8329 AACCTGAGGCTTTCACGAGGCTATGACAGGATATTCTGCGCTCTGTGTACCCCCC 8388
Db 8604 AGCTGAGAGCTTTCACGAGGCTATGACAGGATCTCGCCCCCGGAGACCCCA 8663
Oy 8389 AGACCGAGATATGATGTGAGTGTATGATCATCTTGTCTCTCAATGTGTCTGTGCGCTG 8448
Db 8664 CAACAGAAATAGACTTTGAGGCTTAATACATCATGTCTCTCCAGAGTCAAGTCCGAC 8723
Oy 8449 GGCACAAAGCCCGCGCAATCTACTGACCAAGACCCCTACCTCAATCGCCCG 8508
Db 8724 GACGGGCTGAAAGAGGCTTACTACTTAACTTAACTTAACTTAACTTAACTTAACT 8783
Oy 8509 GCTGCGTGGGAAACAGTTTGAACCTCCCTGCAATTCATGCTGGGAAACATCATCCAG 8568
Db 8784 GCGCGTGGAGACACAAAGACACTTCAATCTTCTGCTGAGCAATATCATG 8843
Oy 8569 TACGCCCGCACTATGAGCTGCGATGCTCTGATGACACACTTCTTCTCATTTCTATG 8628
Db 8844 TTTGCCCCCACTGTGGGAGGATGATAGTACCAATTTCTTTAGCGTCCATAT 8903
Oy 8629 GCTCAAGACAGCTTGAACAGAACTTCAACTTTGAATGATGAGAGCGGTACTCCGTG 8688
Db 8904 GCCAGGATCAGCTTGAAGAGGCTTAACTGTGATCTACGAGACCTGCTACTCATTA 8963
Oy 8689 AGTCCCTTGGACCTCCAGCTATTAATGAAGGTTACATGGGCTTACGCTTTTCTG 8748
Db 8964 GAACCACTGATTTACTCTCAATCATTCAAAGACTCAATGCTTCAAGCGATTTTCACTC 9023
Oy 8749 CACACATACACTCCCAAGCACTGACAGGCTGCTTCAAGCCCTCAGAAAACTTGGGCG 8808
Db 9024 CACAGTTACTCTCAGGTGAATCAATAGGCTGGCGCATGCTTCAAGAACTTGGGCTC 9083
Oy 8809 CCACCTCTCAGAGCTGGAAGAGCGCGGCACTGACGTAGAGGCTGCTTCTTCTCCGT 8868
Db 9084 CCGCCTTGTGAGCTTGAAGACACCGGGCCCGAGGCTCGCGCTTGTGTCTGACAG 9143
Oy 8869 GGGGGGAGAGCGCGCTTGGCGGTGATATCTTTCATTTGGCGGTGAAGACCAAGCTC 8928
Db 9144 GAGAGGAGGCTCTCATATGTGGCAAGTACTTTTCACTGGGAGATTAAGAACAAAGCTC 9203
Oy 8929 AAATCACTCATTTGCGGAGAGCGCTCTGGAATTTATCCAGCTGTTCACGCTGCG 8988
Db 9204 AAATCACTCATATAGGCGCGCTGCGCTGAGACTTGTTCGAGGTTCAGCGGCTGCG 9263
Oy 8989 GCGGCGGGGCGACATTTATCAAGGCTGTGCGGTGCCGACCCGCTTATTTCTTTT 9048
Db 9264 TACAGCGGGGAGACATTTATCAAGCGTGTCTCAATGCCGCGCGCTGCTGTCTGTTT 9323
Oy 9049 GCGCTACTCTACTTTTGTAGGGGTAGGCTTTTCTTACTTCCCGCTG 9098
Db 9324 TGCTACTCTGTCTGCTGCGGAGGGGTAGGATCTACTCTTCCCAACCG 9373

RESULT 8

US-09-034-756-5

Sequence 5, Application US/09034756

Patent No. 6392028

GENERAL INFORMATION:

APPLICANT: RICE, CHARLES et al.

TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C VIRUS (HCV) AND USES THEREOF

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: HOWELL & HAFERKAMP, L.C.

STREET: 7733 FORSYTH BLVD., SUITE 1400

CITY: ST. LOUIS

STATE: MO

COUNTRY: USA

ZIP: 63105

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/034,756

FILING DATE: 04-May-1998

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: HOLLAND, DONALD R.

REGISTRATION NUMBER: 35,197

REFERENCE/DOCKET NUMBER: 6029-4831

TELECOMMUNICATION INFORMATION:

TELEPHONE: 314-727-5188

TELEFAX: 314-727-6092

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 12980 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-034-756-5

Query Match

Best Local Similarity 66.4%; Score 4129.6; DB 3; Length 12980;

Matches 6051; Conservative 0; Mismatches 2969; Indels 90; Gaps 5;

QY 1 ATGAGCAAAATCTTAACCTCAAGAAAGAAACCAAAACCAACCAACCGTCCGCCA 60
DB 342 ATAGGACGAATCTTAACCTCAAGAAAGAAACCAAAACCAACCAACCGTCCGCCA 401
QY 61 GACGTTAAGTTTCGCGGCGCGGCGAGATGTTGGCGAGATATCTTGTGGCGGAGG 120
DB 402 GAGCTCAAGTTCCCGGCGGCGGCGGCGAGATGTTGGTGAAGTTTACTTGTGGCGGAGG 461
QY 121 GGCCTCAAGTTGGGTGTGCGCGGCAAGAAAGACTTTCGAGCGGTCCTCAAGCACTGA 180
DB 462 GGCCTCAAGTTGGGTGTGCGCGGCAAGAAAGACTTTCGAGCGGTCCTCAAGCACTGA 521
QY 181 AGCGCGCAAGCCATCTTAAGATCGCGCTCCACTGCGCAATCTTGGGAAAAACCGA 240
DB 522 AGAGCTAGGCTTATCCCAAGGCACTGCGCGGCGGAGAGCTTGGGCTCAAGCCGGG 581
QY 241 TACCCCTGCGGCTATACGGAATGAGGACTCGGCTGGCGAGGATGGCTCTGTCCCGC 300
DB 582 TACCTTGGGCGGCTCTATGCGAATGAGGAGTTGCGGCGGAGTGGCTCTGTCTCC 641
QY 301 CGAGGTTCCGCTCTTGGGCGGCGGCAATGAGCCCGGCGATAGTGGCGCAACGTGGT 360
DB 642 CGTGGCTCTGCGGCTAGCTGGGCGGCGGCAAGACCCCGGCGATAGTGGCGCAATTTGGT 701

QY 361 AAGTCATGATACCTTAACGTGCGCTTGGCGACCTCATGGGGTACATCCCTGTG 420
DB 702 AAGTCATGATACCTTAACGTGCGCTTGGCGACCTCATGGGGTACATCCCTGTG 761
QY 421 GCGCGCCCGGCTCGCGGCGGCGGCGAGCTTCGCGCATGGGCGTGAAGTCTGGAGAC 480
DB 762 GCGCGCCCGGCTCGCGGCGGCGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 821
QY 481 GGGGTTAATTTGGCAAGGAACTTACCGGTTGCTCTTTTATCTTCTTCTGCGC 540
DB 822 GGGGTTAATTTGGCAAGGAACTTACCGGTTGCTCTTTTATCTTCTTCTGCGC 881
QY 541 CTGCTGCTGATACCAACCGGCTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
DB 882 CTGCTGCTGATACCAACCGGCTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 941
QY 601 TACATGCTGATACCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
DB 942 TACATGCTGATACCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1001
QY 661 CTCACGCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
DB 1002 CTCACGCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1061
QY 721 CCGGCTCAACGGAATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
DB 1062 CCGGCTCAACGGAATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1121
QY 781 CACATGCAATGGTGTGATGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
DB 1122 CACATGCAATGGTGTGATGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1181
QY 841 TCGGCTGCGGATGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
DB 1182 TCGGCTGCGGATGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1241
QY 901 GTCCAGACGCAATGCTCATACCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
DB 1242 GTCCAGACGCAATGCTCATACCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1301
QY 961 GACATGATGATGATGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
DB 1302 GACATGATGATGATGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1361
QY 1021 CCGGAGCTATTAACATCATTAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
DB 1362 CCGGAGCTATTAACATCATTAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1421
QY 1081 TACTTCTATGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
DB 1422 TACTTCTATGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1481
QY 1141 GTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
DB 1482 GTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1541
QY 1201 AGCTTATGACATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1260
DB 1542 AGCTTATGACATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1601
QY 1261 CAGATGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1320
DB 1602 CAGATGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1661
QY 1321 CTGTTTACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380
DB 1662 CTGTTTACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1721
QY 1381 AGTATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1440
DB 1722 CGGCTTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1781

OY	1441	CCAGAAGATATAGAACCTTATTTGCTGGCACTACCCACCAAGGAGTGTGGCGTCTCC	1500
Db	1782	GAGCA-----ACGCCCTTACTGTGTGACATACCCTCCAGACCTTGTGGATTTGTGCC	1835
OY	1501	GCGAAGACTGTGTGTGGGCCAGTGTACTGTCTTCAACCCCAAGCCAGTGTAGTGGACAG	1560
Db	1836	GCAAAAGACGTGTGTGTGGCCCGGTATATTGCTTCACTCCAGCCCCGTGTGTGTGGAAAG	1895
OY	1561	ACCGACAGGCTTGGAGCGCCCACTTACAGTGTGGGGGAGAAATGAGACAGATGCTTCTTA	1620
Db	1896	ACCGACAGGTGGGGCGCGCTTACCTACAGCTGGGGGTGCAATATGATATAGGATGTCTTGTC	1955
OY	1621	TTGAAACAGCACTCCAGCCACCGCTGGGGTCAATGTTCCTGGCTGCAGTGTGAATCTTCT	1680
Db	1956	CTTAAACAACACAGGCCACCGCTGGGCAATTTGGTTCGTGTGTACTGTGATGMACTCACT	2015
OY	1681	GGCTAACACCAAGACTTGGGGGGGCACACCTCGCGTACTAGAGCTGACTTCAAGCCAGC	1740
Db	2016	GGATTTCACCAAAATGTGTGGAGCGCCCTTGTGTCACTGGAGGGGGTGTGGCAACAAC--	2073
OY	1741	ACGGAAGCTGTGTGCCCCACGGAAGTGTTTAGAAAGACATCCTGATACCACTTACCTGAAA	1800
Db	2074	----CCTTGTCTGTGGCCCACTGATTTGTTTCCGAAAGACATCCGAAAGCACTACTCTCGG	2129
OY	1801	TGCGGCTTGGGGCCCTGGCTCAACGCACAAAGGTGCTGATTCGACTACCCCTAAGGCTCTGG	1860
Db	2130	TGCGGCTCCGGTCCCTGTGATTACACCCAGGGTGCATGTGTGACTCCGTAATAGGCTTTGG	2189
OY	1861	CATTACCCCTGCACAGTTAATCAATATACATCTTCAAAAATAGAGATGTATGTGGAGGGGTT	1920
Db	2190	CACATTCCTTGTACCATCATATTAACCATATTCAAAATGACAGATGTAGTGTGGAGGGGCTC	2249
OY	1921	GAGCACAAGCTCAAGGCTGCATGCAATTTCACTCGTGGGGATCGTTTGCACTTGGAGAGAC	1980
Db	2250	GAGCACAAGCTGGAAGGGCGCTGCACAATGCAGCGGGGGCAAGCCTGTGATCTGGAAAGAC	2309
OY	1981	AGAGACAGAATGCAACTGTCTCTCTTTGTTGCACTCCACACAGGATGGGCCATTTTACT	2040
Db	2310	AGGGACAAGTCCGAGCTCAGGCCAATTTGCTGTCTCCACACAACAGTGGCAGGTCCTTCCG	2369
OY	2041	TGCTTTTACTGGAGCTGACCGCGCCCTGTTCGACTGGTCTTTCACCTCCACCAAAATCATC	2100
Db	2370	TGTTTTCACGACCTCTGCAGCCTTGTTCACCGGCTCTCATCCATCTCCACAGAACATTT	2429
OY	2101	GTGACGTACAAATTCATGTATATGAGCTTATCACTGCGCTTCACAAAAATACATGTCGCAATG	2160
Db	2430	GTGACGTGCAAGTACTTGTACGGGGTAAAGGTCAAGCAATCGGCTCTGGGCCATTAAAGTGG	2489
OY	2161	GAGTGGGTAAATCTCTTATTTCTGTGCTCTTAAAGCAGCGCAGAGGTTTGCCTGCTTATGG	2220
Db	2490	GAGTACGTGCTTCTCTGTCTCTCTCGCTTGTGACAGACGCGCCGTCGTCTCTGCTTGTGG	2549
OY	2221	ATGCTCATCTTGTATGGGCGAGGCGGAACAGACACTAGAGAAAGCTGTGTATCTTGTGACAGCT	2280
Db	2550	ATGATGTTACTCATATATCCCAAGCGGAGCGGCTTGGAGAACCTTCGTAAATCTCAATGCA	2609
OY	2281	GCGAGCGAGACTGAGCTGCAAATGAGCTCTCTATATTTTGTATCTTTTGTGGCTGTCTGG	2340
Db	2610	GCAATCCCTGGCCGGGACCGCATCGGTCTTGTGTCTCTCTGTGTCTTGTGCTTGTGGCGTGG	2669
OY	2341	TACATCAAGAGGTGGGGTATGCTCCCTTAACTATATCTCTTCACTGSCCTGTGTCTTT	2400
Db	2670	TATCTGAAGGGTATAGGTGTGGGTGCCCCGAGCGGTCTACGCGCTTCTAAGGATGTGGCTCTC	2729
OY	2401	AGCTTACTGTCTCTAGCATTTGGCCCAACAGGCTTATGCTTATAGACGATCTGTGTATGGC	2460
Db	2730	CTCCGCTCTCTGTGTGGCGGTGCTTACAGCGGGATATGCACTGGAACGGAAGTGTGCGCG	2789
OY	2461	CAGATAGAGCGGCTCTGTGTGTATATGATCACTCTTAACTCTCAACCCCGGGATATAG	2520
Db	2790	TGTGTGTGGCGGTGTGTTCTTGTGTGGGTAAATGGGCGTGACTCTGTCCGCATATTACAA	2849
OY	2521	ACCTTCTCAGCGGGTTTTTGTGTGTGTGTGTGTATCTTCTGACCTGTGGGGAAGCTATG	2580

Db	2550	CGTACATCAAGCTGTGTGATGTGTGGCTTCAGATATTTTCTGACCAAGATAGAAAGCCCA	2909
Qy	2581	GTCCAGAGATGGGACCAACCTATGACAGTGCACGGTGGCCGTGATGGCATCATATGGACC	2640
Db	2910	CTGCACGATGTGGTGTTCCTCCCTCCCAACGTCACGGGGGGGGCGCGATGCCGTATCTTACTC	2969
Qy	2641	GTGCCCATATTCTAACCCAGAGTGTGTGTGGACATAACCAAGTGGCTTTGGCGGTGCTT	2700
Db	2970	ATGTGTGTGTGACACCCGACTGTGTATTGTGACATCAACAACTATCTGGGCATCTTTC	3029
Qy	2701	GGGCTCGCTTACCTCCCTAAAGAGTCTTGTGACGGCGTGCCTGACTCTGTCAAGGACTCAC	2760
Db	3030	GGACCCCTTGGATTTCTTCAAGCCAGTTTGTCTTAAAGTCCCTTACTTGTGCGGTTTCA	3089
Qy	2761	GCTCTACTGAGAGATGTGCACCATGACAAAGCATCTCGCGGGGGGACAGTACGTCCAGATG	2820
Db	3090	GGCTTCTCCGGATCTGTGGCGGCTAGCGCGGAAGATAGCCGGAGGTCAATTACGTCAATG	3149
Qy	2821	GGCGTACTAGCCCTTGGCAGGTGGACTGGACCTTACATCTATGACCACTCACCCCTATG	2880
Db	3150	GCCATCATCAAGTTATGGGGCGCTTACTGGCCACTATGTGTATACATCTCACCCCTCTT	3209
Qy	2881	TCCGATTGGGCTGTAGTGGCTGACGGGACCTGGCGGGCGGCTTGAAGCTATCATCTTC	2940
Db	3210	CGAAGCTGGGGGCAACAGGCTGCGAGATTTGGACCGTGGCTGTGGAAACCAAGTGTCTTC	3259
Qy	2941	AGTCCGATGAGAAAGAAATGATTTGTCTGGGAGACGGACAGCTGTCTTGTGGGGAATT	3000
Db	3270	TCCGAGATGAGAGACCAAGCTCATCAAGTGGGGGCAATACCGCGGTGGGTGATC	3329
Qy	3001	TTAACACGACTTCCCGTGTCCGCCGACTTGTGTGGAGGTCTCTCTTGGCCCACTGAT	3060
Db	3330	ATCAACGGCTTGGCCGCTCTCGCCGTAAGGGGCGAGAGATCTGTTGGCCACGGCAC	3389
Qy	3061	GGCTATACCTCCAAAGGGGTGGAGTCTTCGGCCCCCATCACTGTCTTACGGCCACAGACA	3120
Db	3390	GGAAATGCTTCAAGGGGTGGAGTGTGTGGCGCCATCAAGGGGTAGGCCACAGACAGC	3449
Qy	3121	CGTGGCCCTTTTGGGACCAATAGTGTGAGCATGACAGGGGCGGACAAAGACAAGCT	3180
Db	3450	AGAGGCTCTCCAGGTGTATTAATCACACGCTGTAGCTGGCGGGACAATAACCAAGTGGAG	3509
Qy	3181	GGGGAATTCAGGTCTGTGCCAAGTCACTCAATGCTCTTCTCGAGACATCCATCTGGGG	3240
Db	3510	GGTAGGGTCCAGATCGTGTCAACTGCTACCCAAACTTTCCTGGGCAACGTGCATCAATGGG	3569
Qy	3241	GTTTGTGTGAGACTGTCTACCATGAGAGTGGCAACAAGTCTGGCGGTACACGGGGTCCG	3300
Db	3570	GTAATGTGAGATGTCTTACACAGGGGCGGAAACGAGAACATCGATCACTCCAAAGGTTCT	3629
Qy	3301	GTCACGACAGATGTACTCCAGTGTCTGAGGGGAGCTTATGATGGGTGGCCCAAGCCCCCTGGG	3360
Db	3630	GTCATCCAGATGTATACCAATGTGGACCAAAACCTTGTGGGCTGGCCGCTCTCTAAGT	3689
Qy	3361	ACTAATCTTTGAGCCGTGACCGTGTGAGACCGGTGACCTGTACTGTGTACACGGGAAC	3420
Db	3690	TCCGCTCATTTGACACCTGTGACCTGGCGGCTCTGTGACCTTTACTGTGTACAGAGCAC	3749
Qy	3421	GCTGATGTCAATCCGGGCTCGAAGACCGGGGAGCAAAACGGGAGAGGCTACTTCCCGGAGA	3480
Db	3750	GCCGATGTCAATCCGTGACCGCGGGGAGAGTATGACAGAGGTATGCTTTTCGCCCGG	3809
Qy	3481	CCTCTTTCACACTTGAAGGGGTCTCAGAGAGCCCGGGTGTATGCCCCCAGGGGACCACT	3540
Db	3810	CCCATTTCTTACTTAAAGGCTCTCTGGGGGGGTCCGTGTGTGTGCCCCCGGGACACCC	3866
Qy	3541	GTCCGAGTCTTCCGGGACAGTGTGTGCTCTCGGGCGGTGAGTCAATGATGATTTCAATC	3600
Db	3870	GTGGGCTTATTCAGGGCGCGGGTGTGACACCGTGGAGTGGCTTAAGGCGGTGACTTTATTC	3929
Qy	3601	CCCGTTGACACTGACATGTCTACCGGGTCTCCCACTTTATGTGACAAACAGCACACA	3666

3930 CTTGTGAGAACTAGAGAACCAACCATGATCCCCGGTGTTCAGGACAACTCCCTCCCA 3989
3961 CTTGCTGTGCCCCGAGACTATCAGGTGGGGTACTTGGATGCCCCGACGGAGGGAAG 3720
3990 CAGAGATGCCCCGAGAGCTTCAGGTGGCCCCACCTGGATGCTCCACCGGAGGGTAG 4049
3721 AGCAACAAGTCTCTGTGCAATATGCTGTCAGGGGATATAAGTGTAGTCTTAATCC 3780
4050 AGCAACAAGTCTCTGTGCAATATGCTGTCAGGGGATATAAGTGTAGTCTTAATCC 4109
3781 TCACTGCTGTCACCTCTGGGGTGGGGGCTTCTTCTTAAGGCAATGGCAATATCC 3840
4110 TCTGTGCTGCAACGCTGGGCTTGTGTCTTACATGTCGAAGGCCCATGGGGTGTGATCT 4169
3841 AATATTGAGCTGAGTCAAGATCTGTGACCAACCGGGGCGCCCATCAGTACTCCACATAT 3900
4170 AATATCGAGACCGGGGAGGAACAATTACATGAGCCGACCCCATCAGTACTCCACATAT 4229
3901 GGCMAATTCCTGCGCGATGGGGGCTGTGCGGGCGGCGCTTACGACATCATATATGTGAT 3960
4230 GGCMAATTCCTGCGCGATGGGGGCTGTGCGGGCGGCGCTTACGACATCATATATGTGAT 4289
3961 GAATGCCATCCGCTGACCTTACCAACATCTTGGGATCGGAACGATCTTGTATCAAGCA 4020
4290 GAGTGCCATCTCAACGATGCAATCATCTTGGGATCGGCACTGTCTTGAACCAAGCA 4349
4021 GAGACGCTGGGGTCAACATCATCTGTGCTGCTTACACTTACGCCCCCTGGGTGATGACA 4080
4350 GAGACGCTGGGGGAGAGATGTTGTCTGCGCACTCTACCCCTCGGGCTCGTACT 4409
4081 ACCCCCAACCCCAATAGAGAGAGTGGCCCTTGGGAGAGAGGCGAGATCCCTTCTAT 4140
4410 GTGTCCATCTTACATCGAGAGGTGTCTGTGCAACACCGGAGAGATCCCTTCTTAC 4469
4141 GGGAGGGCGAATCCCTGTCTTATCATCAAGGAGGAAGACATCTGATCTTCTGCACTCA 4200
4470 GGGAGGGCGAATCCCTGTCTTATCATCAAGGAGGAAGACATCTGATCTTCTGCACTCA 4529
4201 AAGAAAAGTGTACGAGCTGCGGCGGCGCCCTTGGGGATATGGCTTGAATCTAGTGGCA 4280
4530 AAGAAAAGTGTACGAGCTGCGGCGGCGCCCTTGGGGATATGGCTTGAATCTAGTGGCA 4589
4261 TACTACAGAGGGTGTGAGCTCTCCGTAATACCACTCAGGGAGAGCTAGTGTGTGCGC 4320
4590 TACTACAGAGGGTGTGAGCTCTCCGTAATACCACTCAGGGAGAGCTAGTGTGTGCGC 4469
4321 ACCGAGCCCTCATGACAGGGATATCTGAGGACTTTTGAATCTCGTATGCACTGCAAGTA 4380
4650 ACCGAGCCCTCATGACAGGGATATCTGAGGACTTTTGAATCTCGTATGCACTGCAAGTA 4709
4381 GCGGTCACTCAAGTGTGAGCTTCAAGTTTGAATCCCACTTACCATTAACCAACAGATT 4440
4710 TGTGTCACTCAAGAGTGTGAGCTTCAAGTTTGAATCCCACTTACCATTAACCAACAG 4769
4441 GTCCTCAAGAGCTGTCTCAGAGTACGAGCGCGGGGTGCGACGCGGTAGGGGAAGTGTG 4500
4770 CTCCCCAGAGATGTGTCTCAGAGTACGAGCGCGGGGTGCGACGCGGTAGGGGAAGTGTG 4829
4501 GGCATTTATAGTATGTTTCACTGTGAGCGAGCTCAGGAATGTTTGAACAGTATAGT 4560
4830 GGCATTTATAGTATGTTTCACTGTGAGCGAGCTCAGGAATGTTTGAACAGTATAGT 4889
4561 CTCTGTAGAGTGTACGAGCGAGCGGGCGGCAATGTGATGCAACCATCGAGAACACC 4620
4890 CTCTGTAGAGTGTACGAGCGAGCGGGCGGCTGTGTGTATAGTCAACGCGCGCGAATCA 4949
4621 GTGAGGCTCAGAGCGATTTTCAACGCGCGGTTTGTGTGTGTGCAAGCAATCTTGTAG 4680
4950 GTTATGAGTGTACGAGCGATTTTCAACGCGCGGTTTGTGTGTGTGCAAGCAATCTTGTAG 5009
4681 TTTTGGAGGCGATTTTCAACGCGCGCTTCAACCATATAGTCCCACTTCTTATCCAGACA 4740
5010 TTTTGGAGGCGCTTTTCAACGCGCGCTTCAACCATATAGTCCCACTTCTTATCCAGACA 5069

4741 AAGCAATCGGGGGAATTTTGTGATCTTAAACGCTTACAGAGGCTACAGTGTGCGCTAGG 4800
5070 AAGCAATCGGGGGAATTTTGTGATCTTAAACGCTTACAGAGGCTACAGTGTGCGCTAGG 5129
4801 GCCAAGCCCCCCCCCTGCTGAGAGCTGATGTGAAGTGTGTGATCTTCAACCTCAAGCC 4860
5130 GCTCAAGCCCCCTCCCATGTGTGGACCAATGTGGAAGTGTGTGATCCGCTTAAACCC 5189
4861 AACTGCTGGGCCCCACACTCTCTGTGTACCGCTTGGGCTGTGTGTGATCAACAGAGTCA 4920
5190 ACCCTCATGGGCGCAACCCCTGTATTAACAGCTGGGCGCTGTGTGATGGAAGTCAAC 5249
4921 CTCACATACCCGGAAGGAATATCATGCGCACTGCAATGCAAGCGGCACTTGTAGAGTATG 4980
5250 CTGACGCAACCAATATCAAT 5309
4981 ACCGACATAGGATCTTGTGAGAGGAGAGTCTTGGCGGCGCTGCGCGATATGCTGCGG 5040
5310 ACCGACATAGGATCTTGTGAGAGGAGAGTCTTGGCGGCGCTGCGCGATATGCTGCGG 5369
5041 ACCGAGTGTGTGTGATCATATGCGCGCTTGCACATTAACAGAGAGCGCTTGTGCGCG 5100
5370 ACAGGCTGCTGTGATATGAGGAGAGATGCTTGTGCTGCGGGAAGCGGCAATTAATCT 5429
5101 GACAGAGAGGCTCTTATGAGAGGAGTGTGATGAGATGAGAGGAGTGTGCTTAAAGCGGCT 5160
5430 GACAGAGAGGCTCTTATGAGAGGAGTGTGATGAGATGAGAGGAGTGTGCTTAAAGCGGCT 5489
5161 CTCATTGAAAGAGGAGGAGGAGGAGTATGAGAGGAGTATGAGAGGAGTATGAGAGGAGT 5220
5490 TACATGAGAGGAGGAGGAGTATGAGAGGAGTATGAGAGGAGTATGAGAGGAGTATGAG 5549
5221 CAGCAGCTTCCAAACAGCTCAAGATTAACCACTGTGAGGCTTATGAGGAGGAGGAGG 5280
5550 CAGCAGGCTGCTGCGGAGGAGGAGTATGAGAGGAGTATGAGAGGAGTATGAGAGGAG 5609
5281 GTGAAACAATTTCTGGGCAACACATGTGAGACTTATGAGGAGGAGTATGAGGAGGAG 5340
5610 CTGAGGCTCTTGTGGGCAACACATGTGAGACTTATGAGGAGGAGTATGAGGAGGAG 5669
5341 GGAATATCAACATGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5400
5670 GGCCTGTCAACGCTGCTGAGTAAACCGGATGCTTCAATGATGCTTTTAAAGGAGG 5729
5401 CTCAACAGTCTGCTGTCACAGACACATATCTTCTCAACATTTTGGGGGCTGAGCTA 5460
5730 GTCACAGGCTCACTAACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5789
5461 GCATCCCAATTTGACACACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5520
5790 GTCGCCAGCTGCTGCGGCGGCGGAGTACGAGGCTTGTGAGGCGCTTGAAGTGTG 5849
5521 GTCGCCAGTGTAGGAGTATGAGGAGTATGAGGAGTATGAGGAGTATGAGGAGTATGAG 5580
5850 GCGGCAATGAGGAGGAGTATGAGGAGTATGAGGAGTATGAGGAGTATGAGGAGTATGAG 5909
5581 GCGGCAATTTGGGGGCTCTGTCGATTTCAAGATCATGTCGAGGAGGAGGAGGAGGAG 5640
5910 GCGGAGTGTGAGGAGGAGTCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5969
5641 GAGAGTGTGTCACATTTGAGTGTGAGGAGTATGAGGAGTATGAGGAGTATGAGGAGTAT 5700
5970 GAGAGTGTGTCACATTTGAGTGTGAGGAGTATGAGGAGTATGAGGAGTATGAGGAGTAT 6029
5701 ATCTGCGCGGCAATTTGTCGCGGAGACGTCGAGGAGGAGGAGGAGGAGGAGGAGGAG 5760
6030 GTCGCGGAGCAATTAATGCGCGGAGGAGTGTGCGCGGAGGAGGAGGAGGAGGAGGAG 6089
5761 AATAGACTATTTGCTTGTGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5820
6090 AACCGCTAATATGAGCTTGTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6149

QY 5821 GAGTCGATGCGTGCAGCCGTGTGACCCAACTTGGCTCCCTTACCATTAACGAGCTG 5880
Db 6150 GAGAGCGTGCAGCGCCCGCTCATCTGACAGACCTCACTGTAACCCAGCTC 6209
QY 5881 CTGAGGAAGCTCCAACTGATTAAGTGAAGACTGCCCATTCCTATGCGGCGCTCGTGG 5940
Db 6210 CTGAGGCGACATGCACTGATGATAGCTCGGAGTGTACACTCATCTGCTCGGCTTCGG 6269
QY 5941 CTCCGCGATGTGCGACTGCGTTTGGACATCCCTAAACAGACTTTAAATAATTTGGCTACC 6000
Db 6270 CTAGGGGACATCTGGGACCTGATATGGAAGGTGCTGAGCGACTTTAAGACCTGGCTAAA 6329
QY 6001 TCCAAATTATTTCCCAAAGATGCCCCGCTCTTGTCTCTGCTCAAAAGGGGTACAG 6060
Db 6330 GCCAAGCTCATGCGCAACCTGCTGGATTCCTTTGTGTCTCTCGACGCGGGGTATAGG 6389
QY 6061 GGCCTGTGGCGCGACCTGGATCATGACACAGGCTGCTTGGCGGGGCCAATATCTCT 6120
Db 6390 GGGGCTGGCGAGGAGCGGCAATTAATGCACTGCTGCGCACTGTGAGCTGAGATCACT 6449
QY 6121 GGCATATGCTCCGCTTGGGCTCCATGAGATCAACGGGGCCTTAAGACTTGATGATATCTGG 6180
Db 6450 GGACATGTCAAAAACGGGACGATAGAGATCGTGGCTCTAGSACTGAGGAAACATGTGG 6509
QY 6181 CAGGGGACCTTTCTTATCATATTTGTACACGGAGGGCGAGTGGCCGAAACCCGCGCA 6240
Db 6510 AGTGGGACGTTCCCATTAACGCCCTACACACGGGCGCCCTGTACTCCCTTCGCGCG 6569
QY 6241 AACTTTAAGTTCGCATCTGAGAGGTGGCGGCTCAGAGTACGGGAGGTGACCCAGCAC 6300
Db 6570 AACTATAAGTTCGCGCTGTGAGGGGTGTCTGCAAGAGATACGTAAGATTAAGCGGGTGG 6629
QY 6301 GGGTCATACCACTACATTAACAGAGACTCACCTGATTAACCTTGAAGTCCCTGCGCACTA 6360
Db 6630 GGGGACTTCCACTAGTATCGGGTATGACTACTGACATCTTAATGCGGTGCGAAGTTC 6689
QY 6361 CCTCTCCCGAGTTCTTTTCTCGGTGAGCGGAGTGCAGATTCATAGTTTGGCCCAACA 6420
Db 6690 CCATCGCCCGCAATTTTTCACAGAAATTGAGCGGGGCGCCCTACATAGGTTTGCCTCCCT 6749
QY 6421 CCGAAGCGGTTTTCCGGGATGAGGTCTGTTCTGCGTTGGGCTTAATTCATTTGTGTGTC 6480
Db 6750 TGCAGGCCCTTGTGCGGAGAGGATATTCATTCAGAGTAACTTCACAGATACCCGCTG 6809
QY 6481 GGGTCCGAGCTTCTTTCGCAACCTTGAACCCGACACAGACGTAATTGATTCATGCTAAACA 6540
Db 6810 GGGTTCGCAATTAATCTTTCGAGGCCGGAACCGGACGTAAGCGTGTGTAACATGCTCACT 6869
QY 6541 GATTCATTCATATCAACGCGCGAGACTGCAGCGCGGCGTTTACGGCGGGGCTCACCCCA 6600
Db 6870 GATCCCTCCCATATTAACAGAGAGGCGCGCGGAGAAAGTTGGCGAGAGGGTCAACCCCT 6929
QY 6601 TCCGAGGCAAGTCTCTCGGCGAGCGAGCTATTCGGACCATGCTGCGGACCACTGCAAC 6660
Db 6930 TCTATGCGCAGTCTCTCGGCGAGCGAGCTGTGCTCATCTCTCAAGGCAACTTGCACC 6989
QY 6661 ACCCAGCGCAAGCTATGATGTGACATGTGTGATGCTAACT-----GTTTC 6708
Db 6990 GCGAACATGACTCCCTGACGCGGAGGCTCATAGAGGCTTAACCTCTGTGAGGAGGAGAG 7049
QY 6709 ATGGGGGCGATGTGACTCGATAGAGTGTGGGTCAAAAGTGTCTTGTGACTCTCTC 6768
Db 7050 ATGGCGGCGCAATCAACGAGGTTGAGTCAAGAAACAAAGTGTATTTCTGGAATCTCTTC 7109
QY 6769 GACCAATGTGTGAGAAAGAGGCGACTTGAAGCTTGTGATCACTAGATACATGCTC 6828
Db 7110 GATCCGCTTGTGCGAGAGGAGATAGCGGAGGCTCCGTAACCGCAAAATTCCTGCGG 7169
QY 6829 CCGAAGAGAGGTTCCCAACAGCTTACCGGCTTGGGACCGGCTGATTAACAACCCAG 6888
Db 7170 AAGTCTCGAAGATTCGCCCGGCGCTGCGCGTTTGGGCGCGGCGGACATTAACAACCCCG 7229
QY 6889 CTTGTGAAATGCTGGAAGAGCGCAGATTACCAACCGGCGCACTGTTGCGGGCTGTGCTCTC 6948

Db 7230 CTAGTAGACGCGGAGAAAGCTGTACTAGCAACCACTGTGTTCATGCTGCGCGCTA 7289
QY 6949 CCTCCTCTAGAGAAACCCCGACGCTCTCCCGCAAGGAGCGCGGACAGTGGGCTTAAGT 7008
Db 7290 CCACTTCACAGTTCCTCTGTGTCTCTGCTCTCGGAAAAAGGTAAGGTGTCTCTCACG 7349
QY 7009 GAGACTCCATAGAGATGCTCCCTTCAACAGCTGGCCATTAACTCTTTGGCCAGCCCCC 7068
Db 7350 GAATCAACCTTATCTAGCTGTGGCGGAGCTTGGCCAAAGTTTGTGGCAGCTCTCA 7409
QY 7069 CCAAGGGGATTCAGGCTTTTCCAGGGGGCGGGCTGGCCGATTCGGGCACTAGAGG 7128
Db 7410 ACTTCGGGCAATTAACGGGCGCAATATGCAACCACTCTCTGAGCGCGC----- 7457
QY 7129 CCTCTGATGATGTGGCTCTTTGAGAGACAGTTTCCATCTTCCATGCCCCCTCGAG 7188
Db 7458 CTTTCTGGCTGCGCCCCCGACTCGAGCTGGAAGTCTTATTTCTTCATGCCCCCTCGAG 7517
QY 7189 GGGGAGCTTGGAGATCCAGACTTGAGCCTGAGCAGGTAGAGCCCAACCCCCCCAG 7248
Db 7518 GGGGAGCTGGGGATCCGGA----- 7537
QY 7249 GGGGGGCTGGAGCTCCCGCTCGACTGGGGTCTGTGCTTACTTGTCCAGAGAGAC 7308
Db 7538 -----TCTCAGCGACGGGTCATGTGACAGGTCAGTAGTGGGCGACAG 7583
QY 7309 GACTCGGTGTGTGCTGCTCATGTCTACTCTGAGACGGGGCTCTAATTAATCTCTGT 7368
Db 7584 GAAAGATGTGCTGTGCTGTCAATGTCTTATTTCTGGAACAGGCGCACTGCAACCCGCTG 7643
QY 7369 AGTCCGGAAGAGGAAGTTACCGATTACCCCTTGAAGCACTCCTGTTGCGATATCAC 7428
Db 7644 GTGCGGAGAAACAAAACTGCCCATCAAGCAGCTAGGACAATCGTTGCTACCGCATAC 7703
QY 7429 AACAGGTGATCTGTACCAACAAAGAGGCTCTCACTAAGGCTTAAAGGTAACTTTT 7488
Db 7704 AATCTGTGTATTCACCACTTACGCAAGCTTGGCCAAAGGAGAAAGTCACTTT 7763
QY 7489 GATAGATGCAAGTGTGCACTCTCTACTAGACTCAGTCTTAAGGACATTAAGCTAAGC 7548
Db 7764 GACAGACTGCAATTTCTGACAGACCATTAACAGAGAGTGTCTCAAGAGGTCAAAAGCAG 7823
QY 7549 GCCTCCAAAGTCAACGCAAGCTCTCAACAGAGAGAGGCTTGCAGTTAACCCACCC 7608
Db 7824 GCGTCAAAAGTGAAGCTTAATCTGTATCCGTAGAGGAGCTTGACGCTGAGCGCCCCA 7883
QY 7609 CATTTGCAAGATCTTAATATGTGGTTTGGGGCTTAAGAGTCCGCAAGCTTGTCCGGAGG 7868
Db 7884 CATTCAGCCAAATTCAAAGTTTGTGCTATGGGGCAAAAGACGTCCGTTGCCATGCGCAAGAG 7943
QY 7669 GCGTTAACCAATCAAGTCCGTGTGGAAGACCTCTGGAAGACTCAGAAACACCAATT 7728
Db 7944 GCGGTAGCCCAATCACTCCGTGTGGAAGACCTTCTGGAAGACAGTGTAAACCACTA 8003
QY 7729 CCGAACCAATTAATGCGCAAAATAGAGTGTCTGCTGTGACCCCAACCAAGGGGCGAG 7788
Db 8004 GACATCAATCATATGCGCAAGAGAGTCTTCTGCGTTCAAGCTGAGAGGGGGGCTGT 8063
QY 7789 AAAGCAGCTGCTTATGCTTAACTGACCTGCGGCTCAGGCTTGGCGAAGATGAGCC 7848
Db 8064 AAGCAGCTGCTCATCTGTTTCCCGACCTGGGCTGGCGCTGTGCGAAGAGTGGCC 8123
QY 7849 CTTTATGACATTAACAACAAATCTCTCAGGCGGTGATGGGGCTTCTTATGATTCAG 7908
Db 8124 CTGTACAGCTGTGTAGCAAGCTCCCTGCGGTATGGAAGAGCTCTACGATTCGA 8183
QY 7909 TATTCCTCCGCTGACGCGGTAGAGTCTCTTGAAGCATGGCGGAAAGAGACCT 7968
Db 8184 TACTACACAGAGAGGGGGTGAATTCCTGTGCAAGCGTGAAGTCAAGAGAACCCCG 8243
QY 7969 ATGGGTTTGTGTATGATACCGGATGTTGACTCAACGCTCACTGAGAGAGCATCAGG 8028

Db 8244 ATGGGGTTCTGATGATACCCGCTGTTTGAATCCACAGTCACTGAGAGGACATCCGT 8303
 QY 8029 ACTGAGAGATCCATATATGGGGCTGCTCTTGCCCGAGAGGCGCCACATGCCATAC 8088
 Db 8304 ACCGAGAGGCAATTTACCAATGTTGACCTGGACCCCAAGCCGCGGCGCATAG 8363
 QY 8089 TCGCTAACTAGAGACTTTTACGTGGAGGGCTTATGTTCAACAGCAAGGGCAAACTGCG 8148
 Db 8364 TCCTCACTAGAGGCTTTATGTTGGGGGCGCTTTACCAATTTCAAGGGGAAATCTCG 8423
 QY 8149 GGGTACAGCGCTTCCGCGCCAGCGGGGTGTCACCACTAGCATGGGAAACCATCACA 8208
 Db 8424 GGGTACCGCAGGTGCGCGGAGCGGCTACTGCAACTAGCTGTGTAACACCTCACT 8483
 QY 8209 TGGTACGTGAAGCTTAAAGGCTTGAAGCTGCAAGGATATCGGCGCCCAAAATCTG 8268
 Db 8484 TGGTCACTCAAGGCGCGGAGCGGCTGTCAGAGCGCAAGGCTCCAGAGCTCCCATCTC 8543
 QY 8269 GTATGCGGCGATGACTTGTGTCATCTCAGAAAAGCGAGGAGACGAGAGCG 8328
 Db 8544 GTGTGTGCGCAGCACTTATGTTATCTGTGAAAGTCCGGGGTCCAGAGAGCGCGCG 8603
 QY 8329 AACCTGAGAGCTTTCAGAGAGCTATGACCAAGTATTTCTGCTGCTGATGACCCGCC 8388
 Db 8604 AGCTGAGAGCTTTCAGAGAGCTATGACCAAGTATTTCTGCTGCTGATGACCCGCC 8663
 QY 8389 AGACCGAGATATCTGAGAGCTATGACCAATTTCTGCTGCTGATGACCCGCC 8448
 Db 8664 CAACCAAGATATGACTTGAAGCTTATGACCAATTTCTGCTGCTGATGACCCGCC 8723
 QY 8449 GGGCCACAAAGCGCGCGCAGATATCTGATACCAAGAGAGCTTACCACTCCCAATGCGCGG 8508
 Db 8724 GAGGCGCGCTGAAAAGAGGCTTACTACTTACCTTACCCGTGACCCCTTGCAGAG 8783
 QY 8509 GCTGCTGAGGAAACATTTAGACACTCCCTGTCAATTCATGCTGAGGAAACATTCATCAG 8568
 Db 8784 GCGCGGAGGAGACAGCAAGACACTCCAGTCAATTCCTGCTGAGGAAACATTCATCAG 8843
 QY 8569 TACGCGCGGACATATGGGCTCGATATGCTGATGACCAATTTCTGCTGATGACCCGCC 8628
 Db 8844 TTTGCCCCCACTGTGGGAGAGATATGATGACCAATTTCTGCTGATGACCCGCC 8903
 QY 8629 GCTCAAGACAGCGCTGACAGAACTCAACTTTGAGATGATGACGAGGAGTATCCGCTG 8688
 Db 8904 GCCAGGATGATGATGACAGGCTTTAATCTGTGAGATGATGACGAGGCTTACTCATATA 8963
 QY 8689 AGTCCCTTGGACCTCCAGCTATATTAATGAAGTTACATGAGGCTTGAAGCTTTTCTG 8748
 Db 8964 GAACCACTGATCTACTCTCAATCATTTCAAAAGACTCATGCGCTCAGCGATTTTCACTC 9023
 QY 8749 CACACATACATCTCCCAAGCAATGACACGCGGTGCTTCAAGCTTCAAGAAATTTGGGCG 8808
 Db 9024 CACAGTTACTCTCAAGTGAATCAATAGGATGCGCCAGTCCCAAGAAATTTGGGCGT 9083
 QY 8809 CCAACCTCTGAGAGGAGAGAGAGCGGCGAGCTGACAGGAGGCTTCAATCTCCGCT 8868
 Db 9084 CCGCTTGTGAGAGCTTGAAGACACCGGCGCGAGAGCTCCGCGTGAAGCTTGTCTCAGA 9143
 QY 8869 GGGGAGAGAGCGGCGGCTTGTGCGGTGATATCTTCAATTTGGGCGAGTGAACCAAGCTC 8928
 Db 9144 GAGAGGAGGCGCTGATATGATGAGCAATCTTCAATGAGGAGTGAACCAAGCAATCTC 9203
 QY 8929 AAATCTCACTCTTATGCGGAGAGCGGCTTGTGATTTATCAGAGTGTTCACCGTGGCG 8988
 Db 9204 AAATCTCACTCAATAGGCGCGCTGAGCTGATCTGTGCTGAGTTCACGCGTGGCG 9263
 QY 8989 GCGGCGGAGGAGCAATTTATCAAGAGGCTGCGGTGCGGAGCGCGCTTATTTGCTTTT 9048
 Db 9264 TACAGCGGAGAGCAATTTATCAAGAGGCTGCGGTGCGGAGCGCGCTTATTTGCTTTT 9223
 QY 9049 GGCCTACTCTTCTTTTGTAGAGGAGTGGCTTTTCTTACTCCCGCTCG 9098
 Db 9324 TGGCTACTCTGCTGCTGAGAGGAGTGGAGATCACTCTCCCAACCG 9373

RESULT 9
 US-08-811-566-1
 ; Sequence 1, Application US/08811566
 ; Patent No. 6127116
 ; GENERAL INFORMATION:
 ; APPLICANT: Rice, Charles et al.
 ; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: David A. Jackson, Esq.
 ; STREET: 411 Hackensack Ave, Continental Plaza, 4th
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/811,566
 ; FILING DATE: 03-MAR-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Jackson Esq., David A.
 ; REGISTRATION NUMBER: 26,742
 ; REFERENCE/DOCKET NUMBER: 1113-1-006
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-487-5800
 ; TELEFAX: 201-343-1684
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 9646 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHEetical: NO
 ; ANTI-SENSE: NO
 ; US-08-811-566-1
 Query Match 45.4%; Score 4128; DB 3; Length 9646;
 Best Local Similarity 66.4%; Pred. No. 0;
 Matches 6050; Conservative 0; Mismatches 2970; Indels 90; Gaps 5;
 QY 1 ATGAGCACAATCTCTAACTCAAGAAACCAAAAGAAACCAACCGTCCGCCACAA 60
 Db 342 ATGAGCACAATCTCTAACTCAAGAAACCAAAAGAAACCAACCGTCCGCCACAG 401
 QY 61 GACGTAAGTTTCCGGGCGGCGGCGCAGATGTTGGCGAGTATATCTTGGCGGAG 120
 Db 402 GACGTAAGTTTCCGGGCGGCGGCGCAGATGTTGGCGAGTATATCTTGGCGGAG 461
 QY 121 GGGCCAGAGTTGGGTGCGCGCGAGCAAGAGAACTTGGAGCGGTCCAGCCAGTGA 180
 Db 462 GGGCCAGATTTGGGTGCGCGCGAGCAAGAGAACTTGGAGCGGTCCAGCCAGTGA 521
 QY 181 AGGCGCAGAGCCATCTCTTAAGATCGGCGTCCACTGCGAAATCTTGGGAAACCAAGA 240
 Db 522 AGAGCTAGAGCTATCTCCCAAGGAGCAAGTGGCGCGGAGGAGCTTGGAGTCCAGCGG 581
 QY 241 TACCCCTGAGCCCTTATACGGAATGAGGAGCTGCGCTGGGAGAGATGAGCTGCTCC 300
 Db 582 TACCCCTGAGCCCTTATACGGAATGAGGAGCTGCGCTGGGAGAGATGAGCTGCTCC 641
 QY 301 CGAGTTCCTGCTCTTGGGAGCCCAATGACCCCGCATAGTTCGCGCAACGTGAGT 360
 Db 642 CGTGGCTTCGCGCTAGAGTGGGCGCCCAAGACCCCGGAGTGGTTCGCGCAATTTGGGT 701

QY 361 AAGGTCAATCGATACCTTAACGTGCGGCTTTGCGACCTATGGGGTACATCCCTGTGTG 420
Db 702 AAGGTCAATCGATACCTTAACGTGCGGCTTTGCGACCTATGGGGTACATCCCTGTGTG 761
QY 421 GGGCCCCGGCTCGGCGGCTGCGCAAGACTTCGCGCATGGCGTGAAGTCTTGGAGAC 480
Db 762 GGGCCCCCTCTTGGAGGGCTGCGCAAGGCTCTGGCGCATGGCGTTCGCGGCTTGGAGAC 821
QY 481 GGGGTAAATTTGGACAAGGAACTTACCGGTTGCTCTTTCTATCTTTCTGCTGACC 540
Db 822 GGGGTAAATTTGGACAAGGAACTTCTTGTGCTCTTTCTATCTTTCTGCTGACC 881
QY 541 CTGCTCTCTGATCAACCAACCGGCTCTGCGCTGCGAAGTGAAGAACTATGACCGAC 600
Db 882 CTGCTCTCTGATCAACCAACCGGCTCTGCGCTGCGAAGTGAAGAACTATGACCGAC 941
QY 601 TACATGGTGAATTAACGATGCAACCATGACAGCATTAATGACGCTCAGGCTGTGTC 660
Db 942 TACATGGTGAATTAATGGCTTAACCTGAGTATGTTGTAAGAGGCGCGCATGCAATC 1001
QY 661 CTCCACGTCCTCGGAGTGGCTCCCGTGCAGAAAGTGGGAAATGCAATCTCAGTGTGATA 720
Db 1002 CTGCACTACTCCGGGAGTGTGCTCTTGTGCTGCGCAAGGTAACGCTCTGAGGTGTGGGTG 1061
QY 721 CCGGTCTCAGCGAAATGTGGCCGTGCAAGCGCCCGGCGCTCAACGAGGCTTTCGGAACG 780
Db 1062 GCGGTGACCCCGACGCTGGCCACCAAGGACGCAAACTCCCAACAAGCAGCTTCGACGT 1121
QY 781 CACATCGACATGTTGTGATGTCGCCACGCTGTGCTGTGCTCTTACATGTTGGGAGACTTC 840
Db 1122 CATATCGATCTGTCTTGTGCGAAGCGCCCTGTGCTGCGCTCTTACATGTTGGGAGACTTC 1181
QY 841 TGGCGGTGGGAGTGTGCTGCGACGCCAAATGTTCAATGTCGTGCGCGACACACTGTGTT 900
Db 1182 TGGCGGTGTGCTCTTGTGTTGTGTCATCTGTTTACCTTCTCTCCAGGCGCACCTGAGACG 1241
QY 901 GTCCAGAATCGAATTTGCTCCATCTACCCCTGTATACATCACTGAGACAACCGCATGGAC 960
Db 1242 ACGAAGACTGCAATTTGTTTATCTATCCCGCCATTAACGGGTCAATCGATGGACATGG 1301
QY 961 GACATGATGATGAATCTGTGCGCCACGCGCTACATGATCTTGGCGTACCGCATGTGTC 1020
Db 1302 GATATGATGATGAATCTGTGCGCCCTTACGCGACGTTGGTGTGATGCTCAGCTGCTCCGATC 1361
QY 1021 CCGGAGCTATTAATGACATCATTAAGCGGGGCTCATTTGGGGGCTCATTTGGGCTTGGCC 1080
Db 1362 CCACAGCTCATATGACATGATGCTGTGCTCACTGGGAGATCTGCGGAGCATAGCG 1421
QY 1081 TACTTCTCTATGACAGAGAGCGTGGCGAAAGTGTGTCATCTTCTGTTGGCGCGG 1140
Db 1422 TATTTCTCATGTGGGAGACTGGGCGAAAGTCTGTGATGTGCTCTATTTGGCGGC 1481
QY 1141 GTGAGCGGCGGCAACCAATACGTGTTGGGGTCTGTGCGCGGAGACCAACCGGCGGCTCAC 1200
Db 1482 GTGAGCGGCGGAAACCCACGTGCAACCGGGGAAAGTGGCGGCGCACACGCGCTGTGTT 1541
QY 1201 AGCTATTTGACATGGGCGCCCAAGGAGAAATCCAGCTGTTAAACCAATGGGAGCTGG 1260
Db 1542 GGTCTCTTACACGAGGCGCCAGAGAAATCCATGATCAACATCAACCAACGCGAGTTGG 1601
QY 1261 CACATCAACCGGACCGGCTGAACTGCAATGATCTCTTGCACACCGGCTTTATCGGCTT 1320
Db 1602 CACATCAATAGCAGCGCTTGAATGCAATGAAACCTTAAACCGGCTGTGTTAGCAGGG 1661
QY 1321 CTGTTCTTACACCAAGCTTCAATCTGTCAAGATGTCCGGAACGATGTCCGCTGCGGC 1380
Db 1662 CTCTTCTTACACCAAAATTCATCTTCAAGCTGTCTCTGAAGGATTTGGCCAGCTCGGA 1721
QY 1381 AGTATGAGGCGCTTCGCGGTGGAGTGGGCGCTTGAATATGAGGATTAATGTCAACCAT 1440
Db 1722 CGCTTACCGATTTTGGCCAGGGCTGGGGTCTTATCAGTATATGCCAAGGAAAGCGCTTC 1781

QY 1441 CCAGAGATATGAGACCTTATTTGCTGGCACTACCAACAAAGGAGTGGGCTGTCTCC 1500
Db 1782 GACGA-----ACGCGCTTACTGTGGCACTACCTCCAAAGACCTTGTGGCATTTGGCC 1835
QY 1501 GCGAAGACTGTGTGTGGCCAGGTACTGTTTTCAACCCCAAGCCATTTGATGTGGCAG 1560
Db 1836 GCAAAAGCGTGTGTGGCCCGGTATATGCTTCACTCCAGGCCCCGTGTGTGGGAAAGC 1895
QY 1561 ACCGAGGCTTGGAGCGCCCACTTACAGTGGGGGAGAAATGACAGATGTCTTCTTA 1620
Db 1896 ACCGAGGCTTGGAGCGCGCTACTACTACAGTGGGGTGAATATATAGGATGTCTTCGTC 1955
QY 1621 TTGAACAGACTGACCAACCGCTGGGGTATGTTGCTGCAAGTGAATGAACTTCTTCT 1680
Db 1956 CTTAACAACCAACGAGCCACCGCTGGGCAATTTGGTTGATCTGATGAATCACTCACT 2015
QY 1681 GGCTAACCAAGACTTGGCGGCGACCAACCTGCGGTACTAGAGCTGACTTCAACGCGCAG 1740
Db 2016 GGATTCACCAAGTGTGGAGGCGCCCCCTTGTGTATGAGAGGAGTGGGCAACAACA-- 2073
QY 1741 ACGGACCTGTTGTGCCCAACGACTGTTTGAAGACATCTGATACACTTACCTTCAAA 1800
Db 2074 -----CTTGTCTGCCCACTGATTTGTTCCGAGACATCCGGAAGCCATATCTCGG 2129
QY 1801 TCGGCTCTGGGCTCTGCTACCGCAAGGTGCTGATGACTACCTCTACAGGCTCTG 1860
Db 2130 TCGGCTCTGGGCTCTGCTGATTAACCCAGGTGATGATGACTACCTCGTATAGGCTTTGG 2189
QY 1861 CATTACCCCTGCACTTAACTATACATCTTCAAAATGAATGATGTGGAGGGGTT 1920
Db 2190 CACTATCTTGTATCAATTAATTAACCATATTAATTAAGTCAAGATGATCGTGGAGGGTTC 2249
QY 1921 GAGCAGAGCTCAACGCTGATGCAATTTCACTGTGGGAGTGTGCAACTTGGAGAGAC 1980
Db 2250 GAGCAGAGCTGGAAGCGGCTGCACTGGACCGGGGAGAAAGCTGTGATCTGGAAGAC 2309
QY 1981 AGAAGCAAGATCACTGTCTCTTGTGTGACTCCACCAAGGATGGGCACTTTTAACT 2040
Db 2310 AGGAGCAAGTCCAGGCTCAGCCCATTTGCTGTCCACCAACAGTGGCAGGCTCTTCCG 2369
QY 2041 TGTCTTACTCGGACCTGCGCGGCTTGTGATCTGTGCTTCCACTCCACCAAAATC 2100
Db 2370 TGTCTTACTCGGACCTGCGCGGCTTGTGATCTGTGCTTCCACTCCACCAAAATC 2429
QY 2101 GTGAGCAATTAATTAATGATGCTATCACTGTGCGCTCAAAATATACATCTGCCATG 2160
Db 2430 GTGAGCAATTAATTAATGATGCTATCACTGTGCGCTCAAAATATACATCTGCCATG 2489
QY 2161 GAGTGGTAACTCTTATTTCTGTCTTATAGCGAGCGCAAGGTTGCGCTGCTTATG 2220
Db 2490 GAGTGGTAACTCTTATTTCTGTCTTATAGCGAGCGCGGCTGCTGCTGCTTATG 2549
QY 2221 ATGCTCATCTTGTGGGCGGCGGAGGAGCACTAGAGAGTGTGATCTTGCAGCGCT 2280
Db 2550 ATGCTCATCTTGTGGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2609
QY 2281 GCGAGCGCACTAGCTGCAATGCTTCTTATTTTGTCACTTTTTCGTGGCTGTGG 2340
Db 2610 GCAATCCCTGGCGGAGAGCAGCGCTTGTGTCTTCTGCTGTGCTTCTTGTGGCTGG 2669
QY 2341 TACATCAAGGCTGGGATGCTCCCTTATGCTATCTTCTTCTGCTGCTGCTGCTT 2400
Db 2670 TATCTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2729
QY 2401 AGCTTCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460
Db 2730 CTCTGCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2789
QY 2461 CAGATAGAGCGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
Db 2790 TGTGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2849
QY 2521 ACCCTTCTCAGCGGTTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580

Db 2850 CGCTACATCAGCTGTGTCATGTGGCTTCAGTATTTTCTGACCAAGATGAAGCCGA 2909
Qy 2881 GTCCAGAGATGGGACCACTATCAGATGCGCGTGGCCGTGATGGCATATATGAGCC 2640
Db 2910 CTGACATGTGGGTTCCTCCCTCCACGTCGGGGGGGGCGCATGTGCGTCATCTTACTC 2969
Qy 2841 GTGCGCATATTTCTACCCAGGATGGTGTGATTAACCAAGTGGCTTTGGGGGTGCTT 2760
Db 2970 ATGTGTGTGTACACCCGACCTGTGATTTGTATGACATCACCAATCTATCTGGCATCTTC 3029
Qy 2701 GGGCTGTGCTTACCTCTTAAAGGTCTTTGACGCGCGTGGCCGTACTTGTGAGGGCTCAC 2760
Db 3030 GGAACCCCTTTGGATTTCTTCAAGCAAGTTTGTCTTAAAGTCCCTTACTTGTGGGTTTCA 3089
Qy 2761 GCTCTACTGAGATGTGCACATGGCAAGCATCTCGGGGGGCAAGTACGTCCAGATG 2820
Db 3090 GGGCTTCTCGGATCTGCGGCTAGCGCGGAAGATACCGGAGGTCAATTAGTCAATG 3149
Qy 2821 GCGCTACTAGCGCTTGGCAGGTGGACTGGACCTTACATCTATGACCACTCACCCCTATG 2880
Db 3150 GCGATATCAAGTATGAGGGCGCTTATCTGGACCTATGTATTAACATCTCACCCCTCTT 3209
Qy 2881 TCGGATTTGGCTGTAGTGGCTGTGCGGACCTGGCGGTGCGCTTGAAGCTTATCATCTTC 2940
Db 3210 CGAGACTGGCGGACCAAGCGCTGTGAGATCTGGCGGTGTGGAACAGTGTCTTC 3269
Qy 2941 AGTCCAGTGAAGAAAGTCAATTTGTGTGGAGCGGAGACAGCTCTTGTGGAGACTT 3000
Db 3270 TCCGGAATGGAAGCAAGCTCATCATCGTGGGGGCAATACCGCGCGTGGGAGATC 3329
Qy 3001 TTACACGAGACTCCCGGTGCGCCGACCTTGTGTGGAGAGTCCCTTGGCGGACGTGAT 3060
Db 3330 ATCAACGGCTTGCCCGTCTCTGCGGTAGGGGCGAGAGATCTGTTGGCGGACGCGAC 3389
Qy 3061 GGCATATCTTCAAGGGGTGAGTCTTCTGCGCCCACTCACTGTCTTACGCCGACAGACA 3120
Db 3390 GGAATGTCTTCAAGGGGTGAGGTGTGCGCGCCATCAAGCGGTACGCCGACAGACG 3449
Qy 3121 CGTGGCTTTTGGGACCATAGTGTGAGCATGAAGGGGCGGCAAGACAGAAAGACT 3180
Db 3450 AGAGGCTCTCAGGGGTATATATCATCAGCGCTGACTGGCGGGCAAAACCAAGTGAAG 3509
Qy 3181 GGGGAATTTCAAGTCTGTCTCAAGTCACTAGTCTTCTCGGAAATCATCTTCCGGG 3240
Db 3510 GGTAGGTTCAGATCGTGTCAACGTCAACCAACTTCTGTGGAAAGTGCATTAATGGG 3569
Qy 3241 GTTTTGTGAGCTGTCTACATGAGCTGGCAACAAGACTTGGCGGCTCAAGGGGTCCG 3340
Db 3570 GTATGCTGACTGTCTACACAGGGGCGGAAAGAGACATCGATACCCAGAGGTCTT 3629
Qy 3301 GTCAACGAGATGTATCTCAAGTCTGAGGGGACTTATGAGGTGGCCCAAGCCCTTGG 3360
Db 3630 GTCATCCAGATGTATCAATGTGGAACCAAGACTTGTGGGCTGGCCGCTCTCAAGT 3689
Qy 3361 ACTTAATCTTTGAGAGCGGTGCAAGTGTGAGCGGTCACTGTATCTGTGTCAAGCGAAC 3420
Db 3690 TCCGCTCATTTGACACCTGTCACTGTGGCTCTCGGACCTTTATCTGTGTCAAGGAC 3749
Qy 3421 GCTGATGTCAATCCGCGCTCGAAGACGCGGGGACAAAGGGGAGGCTACTTCCCGAGA 3480
Db 3750 GCCGATGTCAATCCCGTCCGCGCGGAGGTGATGACAGGGGTAGCTTGTGCGCCCGG 3809
Qy 3481 CTTCTTTTCACTTTGAAGGGGTCTCAAGAGCCCGGTCTATGCCCCCAAGGGCACGCT 3540
Db 3810 CCCATTTCTTACTTTGAAGGGTCTCTCGGGGGGTCTCGTGTGTGCCCCCGGGGACAGCC 3869
Qy 3541 GTGGAGTCTTTGGGGAGCGTGTGTCTCTCGGGCGTGGCTAAGTCAATGAATTTATC 3600
Db 3870 GTGGGCTATTTCAAGGGCGGGGTGTGACCCGTGTGAGGTGAAGGGGTGTGACTTTATC 3929
Qy 3601 CCGTGTGAGCACTGTGACATGTGACCGCGTCCCGCACTTTAGTGAACAAGACACCA 3660

Db 3930 CTTGTGAGAACTTAGAGCAACATAGATGCCCGGTGTTACAGGACAACTCTCTCCA 3989
Qy 3661 CTGTCTGTGCCCCAGACTTATCAGTGTGGGATCTTGTGATGCCCCGATGTGCAATGGAAG 3720
Db 3990 CCAAGCAATGCCCCAGACTTCCAGTGTGGCCACCTGTATGCTCCACCGGAGGGTGAAG 4049
Qy 3721 AGCAACAAGTTCCTGTGCGATATGTGTCTCAGGGGTATTAAGTCTAGTACTTAATCCC 3780
Db 4050 AGCAACAAGTTCCTGTGCGTGTGACAGGCCAGGGGTCTCAAGGTGTGTGTCTAACCCC 4109
Qy 3781 TCAGTGCTGCAACCTCTGGGTTTGGGCGTACTTGTCTAAGGACATGGCATMAATCCC 3840
Db 4110 TCTGTGTCTCAAGCGTGGCTTGTGTGCTTATCATGTCCAAGGCCCATGTGGGTGTGATCT 4169
Qy 3841 AACATTAGACTGAGTCAAGACTGTGTGACACCGGGGCGCCCATCATGTACTCAATAT 3900
Db 4170 AATATCAGGACCGGGGTGAACAATTAACATGAGACGCCCATCATGTACTCAACCTAC 4229
Qy 3901 GGCMAATTCCTGCGCGATGGGGGTGTGCGGGCGGCGCTTACGACATCATATGTGAT 3960
Db 4230 GGCMAATTCCTTGTGCGACGGCGGGGTGTCTAGAGGTGTCTTATGATTAATTTGTGAC 4289
Qy 3961 GAATGCCATGCGGTGACCTTACCAACCATCTTGGCATCGAAACAGTCTTGTATCAAGCA 4020
Db 4290 GAGTGCCACTTCAAGGATGACATCATCTTGGGATATGGCACTGTCTTGAACAAAGCA 4349
Qy 4021 GAGACAGTGTGGGTCAAGCTTAACTGTGTGGTCTTACAGTACGCCCCCTGGGTCAGTACA 4080
Db 4350 GAGACTGCGGGGGGAGAGCTGTGTGTGTGCTGCACTGTCTACCCCTCGGGCTCGTCACT 4409
Qy 4081 ACCCCCAACCCCAATATAGAGAGGTGAGCCCTTGGGAGAGAGGGCGAGTCCCTTCTAT 4140
Db 4410 GTGTCCATCTTAAATCATGAGAGGTGTGTCTGTCCACCAAGAGAGATCCCTTTTAC 4469
Qy 4141 GGGAGGGCGATTCCTCTGTCTTATCATCAAGAGAGAAAGACTGTGATCTTCTGCAATCA 4200
Db 4470 GGCMAAGCTTATCCCTCGAGGTGATCAAGGGGGAAGACATCATCTTCTGCACTCA 4529
Qy 4201 AAGAAAAGTGTGACGAGCTGTGCGGGCGGCTTCTGGGGTATGGGCTTGAATCTAGTGTGA 4260
Db 4530 AAGAAAAGTGTGACGAGCTGTGCGGGAGCTGTGCAATGTGGGCAATATGCGTGTGCC 4589
Qy 4261 TACTACAGAGGTTTGAAGCTGTCTCGTAAATCAACTCAGGGAAGCGTATGTGTGCC 4320
Db 4590 TACTACGCGGTCTTGAAGT 4469
Qy 4321 ACCGAGCCCTCATGACAGGGTATATGAGGAGCTTTGACTCGGTATGACTGCAAGCTA 4380
Db 4650 ACCGAGCTCTCATGACTGTGCTTTACCGGAGACTGTGATGATGATGATGATGATGATG 4709
Qy 4381 GCGGTCACTCAAGTTGATGACTTCAAGTTTGAACCCCATTTCAACCATACACAGATT 4440
Db 4710 TGTGTCACTCAGACAGTGTGATTTCAAGCTTGAACCTTATCTTTATGAGACAAACAG 4769
Qy 4441 GTCCCTCAAGAGCTGTCTCAGGTAGGCAAGCGCGGGGTGTGCAAGGGTGAAGAGCTG 4500
Db 4770 CTCCCCAGAGTGTGTCTCTCAGAGCTCAAGCGCGGGGCAAGACTGTGAGGGGAAGCCA 4829
Qy 4501 GGCATTTATAGTATGTTTCTCACTGTGAGCGAGCTCAGAGAAATTTTGAACAGTGTAGT 4560
Db 4830 GGCATGTACAGATTTTGTGTGACACGGGGGAAGCGCCCTCGGCAATGTTTGAAGTGTGCT 4889
Qy 4561 CTCTGTGAGTGTCAAGACGAGGGGCGGATGTGTAGGCTTCAACCATCGAGACACAC 4620
Db 4890 CTCTGTGAGTGTATGACGAGGGCTGTGTGTGTATGAGTGTCAAGCGCCGCGAGATCA 4949
Qy 4621 GTCAAGGCTCAGGCGTATTTCAACAGCCCGGTTTGTGTGTGTGTGTGTGTGTGTGTGT 4680
Db 4950 GTTATGCTCAGAGCTTATGAACACCCCGGGGCTTCCGTGTGTGTGTGTGTGTGTGTGT 5009
Qy 4681 TTTTGGAGGCAAGTTTTCACCGGCTTCAACACATATGATGCCACTTCTTTCCCAACA 4740
Db 5010 TTTTGGAGGCGCTTATACGGGCTTCAACATATGATGCCACTTCTTATCCAGACA 5069

QY 4741 AAGCAATCGGGGAAAAATTCGCACTTAAACAGCCTACAGGCTAGAGTGCGCTAGG 4800
DB 5070 AAGCAGATGGGGAGAACTTCTTACCTGTAGCGTAAACAGCCACGGTGTGCGTAAAG 5129
QY 4801 GCCAAAGCCCCCCCCCTCTGTGGAGCTCATGTGGAAGTGTGTTGACTCGACTAAGCCC 4860
DB 5130 GCTAAGCGCCCTCCGCCATCGTGGAGCCAGATGTGGAAGTGTGTAATCCGCTTAAACC 5189
QY 4861 ACACCTGGAGCCCAACACTCTCTGTACCGCTTGGAGCTGTATACCAAGAGTCAAC 4920
DB 5190 ACCCTTCATGGGCCAAACCCCTGTCTATACAGACTGGGCGCTGTTCAGAAATGAAGTACC 5249
QY 4921 CTCACACATCCCGTGAAGAAATACATGCGCACCTGACATGCAAGCCGACTTGAAGTCATG 4980
DB 5250 CTGACGCAACCCAACTACCAATATACATGACATGATGCGCGCACTGGAAGTGCATC 5309
QY 4881 ACCGACATAGGGCTTGTGGCAGGGGAGTCTTGGCGGCGTGTGCGCGGTATGTGCTTGG 5040
DB 5310 ACGAGCACTGGGTGCTCGTGGGGGCTCTGGCTGCTGCGCGGTATGTGCTGTCA 5369
QY 5041 ACCGGGTGTGTTGATCATCATCGGCGGCTTGCAATTAACGAGCGAGCGTGTGGCGCG 5100
DB 5370 ACAGGCTGCGTGTGATAGTGGGAGGATGTCTTGTCCGGAAAGCGGCAATTATACCT 5429
QY 5101 GACAAGAGGTCTCTATGAGGCTTTTGAATGAGATGGAAGATGTGCTTGAAGGCGGCT 5160
DB 5430 GACAGGGAGGTCTCTACACAGAGTTCGATGAGATGGAAGATGCTCTACAGCACTTACCG 5489
QY 5161 CTCATTGAAGAGGGGCGAGCGGATAGCCGAGATGCTGAATGCAAGATTCAGAGCTTATG 5220
DB 5490 TACATCGAGCAAGGAGATGATCTGCTGAGCAGTTCAAGAGAGGCGCTCGGCTCTG 5549
QY 5221 CAGCAAGCTTCAAAACAAGCTCAAGACATACACCACTGTGAGGCTTCAAGGCCAAG 5280
DB 5550 CAGACCGGCTCCCGCAAGCAGAGAGTATACCCCTGTGTCCAGACCACTGGCAGAA 5609
QY 5281 GTAGAACAATTCGTGGGCCAAACACATGTGAACTTCAATTAAGCGGCACTCAATACCTGCA 5340
DB 5610 CTCAGAGTCTTCTGGGCGAAGCACAATGTGAATTTCAATGAGGATACAACTACTTGGCG 5669
QY 5341 GGACTATTAACAATGCGCGAGGAACCTGAGTACTTCAATGATGGGCTTCAAGCCGCC 5400
DB 5670 GGCTGTCAACGCTGCTGTGTACCCGCAATGCTTCAATGATGAGCTTTTACAGCTGCC 5729
QY 5401 CTCACAGATCCGCTGTCAACAAGCACCACTATCTTCAACATTTTGGGGGGCTGAGCTA 5460
DB 5730 GTCAACAGGCCACTAACAATGGGCCAAACCTCTCTTCAACATATTTGGGGGGGTGGGTG 5789
QY 5461 GCATCCCAAAATTGCAACACCCGCGGGGCACTGGCTTCTGTGTCAATGAGCTAGTGGGA 5520
DB 5790 GGTGCCAGCTGCGCGCCCGGGTGCCTACCGGCTTGTGGCGGCTTGTAGCTGCG 5849
QY 5521 GCTGCGGTAGGAGATATAGGCTTATAGTGAAGTGTGAGTGAATCTGTGCGAGGATATG 5580
DB 5850 GCCGCACTGGCAGAGGTGAGCTGGGAAAGTCTCTGTGACATTTTGGCAGGATATGCG 5909
QY 5581 GCGGGCATTTTGGGGGCTCTGTGSCATTTCAAGATCATGTGCGAGAAACCTTCATG 5640
DB 5910 GCGGGCGGTGGGAGGCTCTGTGATCATTTCAAGATCATGACGGTGAAGTCCCTCCACG 5969
QY 5641 GAGATGTGTCATCACTTGTGCTGTGGAATTTGTCTCCGGGTGCTTGTGATGAGGAGTC 5700
DB 5970 GAGGACCTGTGTCAATCTGTGCGCGCACTCTGTGCTGTGAGGACCTTGTAGTGGTGTG 6029
QY 5701 ATCTGCGGGGCATTTCTGCGCGCAACGTGGGACCGGGGAAAGCGCGCTGCATATGATG 5760
DB 6030 GTCTGCGAGCAATATCTGCGCGCAGCTTGGCCCGGCGAGGGGCGCATGTCAATGATG 6089
QY 5761 AATAGACTCATGTGCTTTGCTTCCAGAGAAATCATGTGCGCCCACTACATGATGAG 5820
DB 6090 AACCGGCTAATAGCTTGGCTCTCCGGGGGAAACATGTTTCCCCACGCACTACGTGCG 6149

QY 5821 GAGTCGAATCGCTGCGAGCGTGTGACCAACTATCTTGGCTCCCTTACATTAACAGCCTG 5880
DB 6150 GAGAGCGATGAGCGCCGCCGCTCACTGCATATCTACAGCGCTCACTTAACCCAGCTTC 6209
QY 5881 CTCAGAAAGCTCCAAACTGGAATTACTGAGGACTGCCCCATCCCATGGCGGCTGTGG 5940
DB 6210 CTGAGGCGACTGATATGATGATGATGACTGAGGTGATCACTCAATGCTCCGTTCTGG 6269
QY 5941 CTCGCGATGTGTGGGACTGAGGTGTCACCAATCTTAACAGACTTTTAAATTTGGCTGACC 6000
DB 6270 CTAAAGGACATCTGGGACATGATATGCGAGGTGTGAGGACCTTTAAGACCTGGCTGAAA 6329
QY 6001 TCCAAATTAATCCCAAAAGATGCCCGGCTCCCTTGTCTCTGTCTAAAGGGGTACAAG 6060
DB 6330 GCCAAGCTATGCAACACTGCTGGGATTTCCCTTGTGTCTCTGCGAGCGCGGTATAG 6389
QY 6061 GCGTGTGGGCGGCACTGGCATCATGACCAACAGGTGCTTGTGGGCGGCAATATCTCT 6120
DB 6390 GGGGTCTGGAGAGAGACGGCATTAATGACACTGCTGCACTGTGAGCTGATGATCACT 6449
QY 6121 GGCATATGTCGGCTTGGGCTCCATGAGAAATCACGGGCTTAAGACTGATGAATATCTGG 6180
DB 6450 GCAATGTCAAAACCGAGACGATGAGGATGCTGCTGCTTAAGACCTGCAAGAACTGTGG 6509
QY 6181 CAGGGGACCTTCTATCAATGTTTACAGGAGGCGCAAGTGTGCTGCGAAACCCGCGCA 6240
DB 6510 AGTGGAGCTTCCCAATTAACGCTTACACAGCGGCGCCCTGTATCTCCCTTCTGGCGCG 6569
QY 6241 AACTTAAGTGGCCCATCTGAGAGGTGGGCGCTCAGACTGCGGAGGTGACGCGAC 6300
DB 6570 AACTTAAGTGTGCGCTGTGAGAGGTGTCTGAGAGAAATAGGTGAATTAAGCGGGTGG 6629
QY 6301 GGGTCAATACCATCACTAATAACAGGACTCACCACTGATTAATCTTGAATGCCCTGCAACTA 6360
DB 6630 GGGGACTTCCACTAGTATCGGATATGACTATGACTGACATCAATCTTAATAGCCCGTGCAGATC 6689
QY 6361 CCTCTTCCGAGTCTTCTTCTGAGGTGACGAGTGAATCATAGTGTGCTGCCCCACA 6420
DB 6690 CCATGCCCGGAATTTTCAAGAAATGACAGGGGTGCGCTCATAGTGTGTGGCCCCCT 6749
QY 6421 CCGAAGCGGTTTTCCGGGATGAGGTCTGCTTGTGCGGCTTAATCATTTGTGCTG 6480
DB 6750 TGCAGGCTTGTGCTGCGGAGGAGTATCATTTCAAGTATGACTCAACAGTACCCGGTGG 6809
QY 6481 GGGTCCAGCTTCTTTCGCAACCTGAACCCGACACAGACGTATGATGTCAATGCTAACA 6540
DB 6810 GGGTCCGAATTAATCTTTCGAGCGGCAACCGGAGGTAGCCGTGTGACGTCAATGCTCACT 6869
QY 6541 GATCCATCTCATATACAGCGGAGACTGCAAGCGCGGCTTTAGCGCGGGGTACACCCCA 6600
DB 6870 GATCCCTCCCATTAACAGCAGAGCGCGCGGAGAAAGTGTGCGGAGAGGTACACCCCT 6929
QY 6601 TCCGAGGCAAGCTCTCGGCGAGCGCAGCTATGCGGCACTATGCTGCGAGCACTGACCC 6660
DB 6930 TCTATGCGAGCTCTCGGCGAGCGCAGCTGTCCGCTCACTTCAAGGCACTTGCACCC 6989
QY 6661 ACCCAGGCAAAAGCTATGATGTGCAATGTGATGCTAACT-----GTTT 6708
DB 6990 GCCAACATGACCTCCCTGACCGCGAGCTCATGAGGCTAACCTCTGTGAGAGCAGAGAG 7049
QY 6709 ATGGGGGCGATGTGACTGGAATAGTGTGGGTCCAAAGTGTGCTTCTGACCTCTCTC 6768
DB 7050 ATGGGGGCGCAACATCAACAGGGGTGTGATGAGAAACAAAGTGTGATTTCTGACTCTTCT 7109
QY 6769 GACCCAAATGTGGAAGAAAGAGAGCAGCTTGAAGCTTGCATACCATCAGAAATCATGCTC 6828
DB 7110 GATCCGCTGTGTGCAAGAGAGATGAGCGGAGGCTTCCGTATCCCGCAAAATTTGTGCG 7169
QY 6829 CCCAAGAAAGTGTCCACAGCTTTTACCGGCTGTGGCAACGCGCTGATTAACAACCAACCG 6888
DB 7170 AAGTCTCGAGAAATTTCCCGCGGCGCTGCGGCTTTGGGCGGCGGCACTAACACCCCGG 7229
QY 6889 CTTGTGGAATCGTGAAGAAAGCCAGATTACCAACCGGCACTGTGTGCGGCGTGTGCTCTC 6948

Db 7230 CTAGTAGAGACGTGAGAAAAGCCCTGACTAGAACCACTGTGTGCTCAATGGCTGCCCCCTA 7289
Qy 6949 CTTCTCTCTAGAAAACCCCGAGCGCTTCCCCTCAAGAGCGCCGAGCAAGTGGGCTTAAGT 7008
Db 7290 CCACCTCTCAGGGTCCCTCTGTGTCTCCGCTTCGAAAAAGCTTAAGGTGTCTCAC 7349
Qy 7009 GAGGACTCCATAGAGATGCCCTTCAACAGCTGGCCATTAACTCTTGTGGCAGCCCCC 7068
Db 7350 GAATCAACCTTATCTAGCTGCTTGGCCGAGCTTGCACCAAAAGTTTGGAGCTCTCA 7409
Qy 7069 CCAAGCGCGATTCAAGGCTTTTCCAGCGGGCGGGCGCTGCAGATTCCGGCAGTCAGACG 7128
Db 7410 ACTTCCGGCATTTACGGGCGCAAAATACGAACAATCTCTGAGCCCCG 7457
Qy 7129 CTTCTGATGATGTTGGCCCTTTTGGAGACAGGTTCCATCTTCTTCATGCCCCCTTGAG 7188
Db 7458 CTTCTGAGCTGCCCCCGGACCTCCGAGCTTGAATCTTATCTTCCATGCCCCCTTGAG 7517
Qy 7189 GGGGAGCTTGGAGATCCGAGCTTGGAGCTTGGAGAGGTAGAGCCCCCAACCCCCCAG 7248
Db 7518 GGGGAGCTTGGAGATCCGAGCTTGGAGAGGTAGAGCCCCCAACCCCCCAG 7537
Qy 7249 GGGGAGGCTGAGCTCCGCTCGGACTCGGGGTCTGTCTTACTTGTCCGAGAGAGAC 7308
Db 7538 -----TCTCAGCGACGGGTCAATGTGTCAAGGTAGTGGGGCCGACAG 7583
Qy 7309 GACTCCGTGCTGTGTGCTCTCATATCTCACTCTGACCGGGGCTTAAATCACTCTTGT 7368
Db 7584 GAAGATTCCTGTGTGCTCTCATATCTTATTTCTTGGACAGGCGCACTGTACCCCTGTGC 7643
Qy 7369 AGTCCGGAAGAGGAGATTACCGATTACCCCTTGGCAACTCCCTGTGGCATATCAC 7428
Db 7644 GCTCCGGAAGAACAAATCTGCCATCAAGCACTGAGCAACTGTGTTGCTACGCAATCAC 7703
Qy 7429 AACAGGTGTATCTGTAACCAACAAAGAGCGCTCACTAAGGGCTTAAAGATCTTTT 7488
Db 7704 AATCTGTGTATTTCAACCACTTCAACGAGTGTTCGAAAGGCAAGAAAGTCACTTT 7763
Qy 7489 GATAGATGCAAGTGTCTGACTCTTACTACGACTCACTCTTAAAGACATTAACTGACG 7548
Db 7764 GACAGACTGCAAGTCTGAGCAGCCATTACAGAGCTGTCTCAAGGAGTCAAGAGCGC 7823
Qy 7549 GCTCCAAAGGTCAACGAGCTCTCAACATGAGAGAGGCTTGGCAGTTAAACCCACCC 7608
Db 7824 GCGTCAAAAGGTGAAGGTAACTTGTATCGGTAGAGAAAGCTTGTACGCTGAGCGCCCA 7883
Qy 7609 CATTTGCAAGATCTAAATATGGGTTTGGGCTTAAAGAGGTCCGACCTTGTCCGGAGG 7668
Db 7884 CATTTAGCCAAATCTCAAGTTTGGCTATGGGGCAAAAGACGTTCGTTGCCATGCCAAGAG 7943
Qy 7669 GCGGTTAACCAATCAAGTCCGTGTGAGAGAGACTCTCTGAGGACTCAAGAAACCAAT 7728
Db 7944 GCGGTAGCCCAATCAACTCCGTGTGAGAAAGACTCTCTGAGAGCACTGTAAACCAAT 8003
Qy 7729 CCCAACCACTTATGGCCAAATATGAGGTGTCTGCGTGAACCCCAAGAGGGGGGAG 7788
Db 8004 GACATCACTCAATGAGCAAGAACAGAGTTTCTGCTTCAAGCTTGAAGAGGGGGGCTGT 8063
Qy 7789 AAAGAGCTGCGCTTATGTTTATCCCTGACCTCGGCTCAAGGCTCTGAGAAAGTGGCC 7848
Db 8064 AAAGAGCTGCTGTCACTGCTGTCTCCGACCTGGGCGCTGGCGGTGTGAGAGATGGCC 8123
Qy 7849 CTTTATGACATTAACAAAACTTCTCAGCGGTATGGGGCTTTTATGATTTCCAG 7908
Db 8124 CTGTACACATGTGTAGCAAGCTCCCTGCGGTATGGGAAGCTCTTACGGAATTTCAA 8183
Qy 7909 TATTTCCCGCTGAGCGGTATGAGTTTCTCTTGAAGCATGGGCGAAAAAGAGACCT 7968
Db 8184 TACTCACAGAGACGGGGTTGAATTTCTCTGACAGCTGAGAGTCAAGAAAGACCCCG 8243
Qy 7969 ATGGGTTTTCGATATGATCCGATGCTTTGACTCAACCGTCACTGAGAGAGCATGAG 8028
Db 8244 ATGGGTTTCTGATATGATACCCGCTGTTTGACTCCACAGTCACTGAGAGCATCGT 8303
Qy 8029 ACTGAGAGATTCATATATCGGGCTGTCTTGTCCGAGAGAGCCCACTGCTATACAC 8088
Db 8304 ACGAGAGAGCAATTTATCAATGTGTGACTGAGACCCCAAGCGCCGTGGCATACAG 8363
Qy 8089 TCGCTAACTGAGAGCTTTTACGTGGAGGGCTATGTTTCAACAGACAGGCAAAACCTGC 8148
Db 8364 TCCCTACTAGAGAGCTTTATGTTGGGGGCTCTTACCAATTTCAAGGGGAAATCTGC 8423
Qy 8149 GGGTACAGGGGTTTCCGCGCAGCGGGGTCTCAACCACTAGATGGGGAACCATACACA 8208
Db 8424 GGCCTACCGCAGGTGCGCGGAGCGGGGTACTGACAACTAGTGTGTAAACCTCACT 8483
Qy 8209 TGTACGTGAAGCTTTAGCGGCTTGAAGCTGTGAGAGATTAATCGGCCCAATGCTG 8268
Db 8484 TGTCTACATCAAGGCCCGGAGCGCTGTGCGCGCGAGGGGTCCAGAGCTGCACCATCTC 8543
Qy 8269 GTATGCGCGATGACTTGTGTATCTCAGAAAGCCAGGAGACCGAGAGAGACGCG 8328
Db 8544 GTGTGTGCGACCACTTATGTGTGTATGTGAAAGTGCAGGGGTCCAGAGAGACGCGCG 8603
Qy 8329 AACCTGAGAGCTTCAACGAGGCTATGACCAAGATTTCTGCCCCCTCTGTGTGACCCCC 8388
Db 8604 AGCTGAGAGCTTCAACGAGGCTATGACCAAGATTTCTGCCCCCTCTGTGTGACCCCC 8663
Qy 8389 AGACCGAGATATGATCTGAGAGCTGATTAACATCTTGTCTCTCAAAATGTGTCTGCGCTG 8448
Db 8664 CAACCAAGAAATCACTTGGAGCTTATTAACATCTGCTCTTCAACAGTGTACGTCGACAC 8723
Qy 8449 GGGCCCAAGAGCGCGCAGATATCTACCTGACCAAGAGACCTTCACTTCAATGCGCGG 8508
Db 8724 GACGGCGCTGAAAGAGAGGTCTTACTTACTTACCTTACCCGTGACCTTCAACCCCCCTGCGAGA 8783
Qy 8509 GCTGCTGGGAAAACGTTAGACACTCCCTGTCAATTCATGTGCTGGAAAACATCATCAG 8568
Db 8784 GCGCGTGGGAGACAGACACACATCTCCAGTCAATTCCTGTGCTAGGCAATPATCATG 8843
Qy 8569 TAGCCCCCACTATGGGTGCGCATGTGTCTGATGACACACTTCTTCTCATTTCTCATG 8628
Db 8844 TTTGCCCCCACTGTGGGGAGAGATGATCTATATGCCATTTCTTTTAACTGTCTATTA 8903
Qy 8629 GCTCAAGACAGCGCTGAGCCAGAACTCAACTTTGAGATGATGAGAGCGGTGACTCGGTG 8688
Db 8904 GCCAGGATATAGCTTGAACAGGCTTTTAACTGTGAGATCTACGAGAGCTGTACTCATTA 8963
Qy 8689 AGTCCCTTGGACCTCCAGCTATATTTGAAGGTTACATGGGCTTGAACGCTTTTCTCTG 8748
Db 8964 GAACCACTGATCTTACTCTCAATCTTCAAAAGCTCCATGGCTCAGCGCATTTTCACTC 9023
Qy 8749 CACACATTAACCTTCCACGAACTGACAGGGGTGCTTACGCTTCAGAAAATTTGGGGCG 8808
Db 9024 CACAGTTACTCTCAGGTGAATTAATGAGGTGCGGCAAGCTTCAGAAAATTTGGGGCTC 9083
Qy 8809 CCACCCCTGAGAGCTGGAAGAGACCGGGACGTGAGTGAAGGCGTCTCATCTCCGT 8868
Db 9084 CCGCTTGTGAGACTTGTGAGACACCGGGCCGAGAGGCTCGCGCTGTAGCTTGTCTCAGA 9143
Qy 8869 GGGGGGAGAGCGGCGCTTGTGCGGTGATATCTTCAATTTGGCGGTGAAGACCAAGCTC 8928
Db 9144 GAGGAGAGGGCTGCTCATATGTGCAAGTACTCTTCAACTGGGCAATGAAGAAAGCTC 9203
Qy 8929 AAATCACTCAATTTGCCGGAAGCGGCTCTCTGGAATTAATCCAGCTGTTCACCTGTGCG 8988
Db 9204 AAATCACTCAATTAACGCGCGGTGCGGTGAGTCTGTGCTGTGTGTTCACGCGCTGCG 9263
Qy 8989 GCGGCGGGGAGCAATTTATCAAGAGTGTCCGTCGCGGACCCGCTTATTTGCTCTT 9048
Db 9264 TACAGCGGGGAGACATTTATCAAGGCTGTCTCATGTCCGCGGCCGCTGTCTGTCTT 9323
Qy 9049 GGCCTACTCTTCTTTTGTAGGGGTAGCGCTTTTCTTACTCCCGCTCG 9098
Db 9324 TGCCTACTCTGCTCGCTGAGGGGTAGGATCTACTCTCTCCCAACCG 9373

RESULT 10

US-09-034-756-1

Sequence 1, Application US/09034756

Patent No. 6392028

GENERAL INFORMATION:

APPLICANT: RICE, CHARLES et al.

TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C VIRUS (HCV) AND USES THEREOF

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: HOWELL & HAFERKAMP, L.C.

STREET: 7733 FORSYTH BLVD., SUITE 1400

CITY: ST. LOUIS

STATE: MO

COUNTRY: USA

ZIP: 63105

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/034,756

FILING DATE: 04-May-1998

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: HOLLAND, DONALD R.

REGISTRATION NUMBER: 35,197

REFERENCE/DOCKET NUMBER: 6029-4831

TELECOMMUNICATION INFORMATION:

TELEPHONE: 314-727-6092

TELEFAX: 314-727-5188

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 9646 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-034-756-1

Query Match 45.4%; Score 4128; DB 3; Length 9646;

Best Local Similarity 66.4%; Pred. No. 0; Matches 2970; Indels 90; Gaps 5;

Matches 6050; Conservative 0; Mismatches 2970; Indels 90; Gaps 5;

QY 1 ATGAGCACAATCTTAACTCAAGAAAAAACAAGAAAAACCAACCGTCCGCAAA 60
DB 342 ATGAGCACAATCTTAACTCAAGAAAAAACAAGAAAAACCAACCGTCCGCAAG 401
QY 61 GACGTTAAGTTCCGGGCGCGGCGCAATCGTTGGCGGAGTATCTTGTCCCGCAAG 120
DB 402 GACGTTAAGTTCCGGGCGCGGCGCAATCGTTGGCGGAGTATCTTGTCCCGCAAG 461
QY 121 GGGCCGAGTTGGGCGGCGGCGCAAGAAAGTTCGAGCGGCTCCAGCCAGTGGGA 180
DB 462 GGGCCGAGTTGGGCGGCGGCGCAAGAAAGTTCGAGCGGCTCCAGCCAGTGGGA 521
QY 181 AGGCGCAGCCATCTTAAGATCGGCGCTCACTGGCAATCTTGGGGAACCAAGGA 240
DB 522 AGGCGCAGCCATCTTAAGATCGGCGCTCACTGGCAATCTTGGGGAACCAAGGA 581
QY 241 TACCCCTGGCCCTTATAGGGAATAGGGAATCTGGCTGGGCAAGATGCTCTGCCCC 300
DB 582 TACCCCTGGCCCTTATAGGGAATAGGGAATCTGGCTGGGCAAGATGCTCTGCCCC 641
QY 301 CGAGGTTCCGCTCTTGGGCGGCGGCAATGAGCCCGGCAATAGGTCGGCAAGTGGGT 360
DB 642 CGGCTCTCGGCTAGGCTGGGCGGCGGCAATGAGCCCGGCAATAGGTCGGCAATTTGGGT 701

QY 361 AAGGTCATGATACCTTAAGCTGCGGCTTTGGCGAAGCTCATGAGGATATCCCTGTCTGT 420
DB 702 AAGGTCATGATACCTTAAGCTGCGGCTTTGGCGAAGCTCATGAGGATATCCCTGTCTGT 761
QY 421 GGGGCCCCGCTGGGCGGCGGCTGGCAAGCTCTGCGCAATGGGCTGAGAGATCTTGGAGAG 480
DB 762 GGGGCCCCGCTGGGCGGCGGCTGGCAAGCTCTGCGCAATGGGCTGAGAGATCTTGGAGAG 821
QY 481 GGGGTTAATTTGGCAAGAGAACTTACCCGAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCT 540
DB 822 GGGGTTAATTTGGCAAGAGAACTTACCCGAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCT 881
QY 541 CTGCTGTCTGATCAACCAACCCCGGCTCTCGGCGGAGAGTGAAGATCAATCAATCGGCG 600
DB 882 CTGCTGTCTGATCAACCAACCCCGGCTCTCGGCGGAGAGTGAAGATCAATCAATCGGCG 941
QY 601 TACATGATGATTAACGATGCAACCAATGACAGATTAAGTCAAGCTCCAGGCTGTCTGT 660
DB 942 TACATGATGATTAACGATGCAACCAATGACAGATTAAGTCAAGCTCCAGGCTGTCTGT 1001
QY 661 CTCACAGTCCCGGAGTGTCTCCGAGTGAAGAAAGTGGAGAAATGATCTGAGTGTGATGA 720
DB 1002 CTCACAGTCCCGGAGTGTCTCCGAGTGAAGAAAGTGGAGAAATGATCTGAGTGTGATGA 1061
QY 721 CCGGCTCAACGGAATGAGCCGCTGACGCGGCGGCGGCGGCTTACGCGAGGCTTGGCGAG 780
DB 1062 CCGGCTCAACGGAATGAGCCGCTGACGCGGCGGCGGCGGCTTACGCGAGGCTTGGCGAG 1121
QY 781 CACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 1122 CACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1181
QY 841 TGGGAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 1182 TGGGAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1241
QY 901 GTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 1242 GTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1301
QY 961 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 1302 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1361
QY 1021 CCGGAGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1362 CCGGAGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1421
QY 1081 TACTTCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
DB 1422 TACTTCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1481
QY 1141 GTGAGAGGCGGCAACCAATCTGTTGGGAGTCTGCGGCGGCAAGCAACCGGCGGCTCAC 1200
DB 1482 GTGAGAGGCGGCAACCAATCTGTTGGGAGTCTGCGGCGGCAAGCAACCGGCGGCTCAC 1541
QY 1201 AGCTTATTGACATGAGGCGGCGGCAAGAAATCACTGTTTAAACAATGAGCACTGGA 1260
DB 1542 AGCTTATTGACATGAGGCGGCGGCAAGAAATCACTGTTTAAACAATGAGCACTGGA 1601
QY 1261 CACATCAATGAGCAAGGCTTGAATCAATGAAAGCTTAAACCCGCTGTTTAAAGAGG 1661
DB 1602 CACATCAATGAGCAAGGCTTGAATCAATGAAAGCTTAAACCCGCTGTTTAAAGAGG 1721
QY 1321 CTGTTCTAACCCCAAGCTTGAATCTGTCAGAGATGTCGCGGAAGCAAGTCCGCTGCGCG 1380
DB 1662 CTGTTCTAACCCCAAGCTTGAATCTGTCAGAGATGTCGCGGAAGCAAGTCCGCTGCGCG 1721
QY 1381 AGTATGAGGAGCTTCCGAGTGGAGTGGGAGCTTCAATATAGGATATATGTCACCAAT 1440
DB 1722 AGTATGAGGAGCTTCCGAGTGGAGTGGGAGCTTCAATATAGGATATATGTCACCAAT 1781

OY	1441	CCAGAGATATGAGACCTTATTTGCTGGCACTACCAACAGGACAGTGTGGCGTGGCTCC	1500
Db	1782	GACGA-----ACGGCCCTTACTGTGGCGACTAACCTCCAMACCTTGTGGCATTTGGCCC	1835
OY	1501	GCGAAGA CTGTGTGGGCGCAGTGTACTGTTTACCCCGACCGACAGTGTAGTGGGACG	1566
Db	1836	GCAAGAGCGTGTGTGGCCCGGTATATTGCTTCACTCCAGCCCGTGTGTGTGGAAAG	1899
OY	1561	ACCGACAGGCTTGGAGCGGCCCACTTACACGTGGGGGAGAA TGAGACAGATGTCTCTTA	1620
Db	1896	ACCGACAGGTGGGGGGCGGCTTACCTACAGCTGGGGTGCMAATGATACGGAATGTTCGTTC	1955
OY	1621	TTGAACAGACACTCCACACACCGTGGGGTTCATGTTCCGCTGCACGTGATGAATCTTCT	1680
Db	1956	CTTAACAAACACAGGCGCACCGCTGGGCAATTTGGTTCCGTTTGACTCGATGAATCAACT	2015
OY	1681	GGCTACACCAAGACTTGGGGGGCACACCGCTGCCGTACTAGAGCTTCAAGCGCACG	1740
Db	2016	GGATTACCAAAAGTGTGGAGAGCGCCCTTGTGTCACTGGAGGGGTGGGCAACAAC--	2073
OY	1741	ACGGACCTGTGTGGCCCGACGGAAGCTGTTTATGAGAGCATCCGTGAATCACTTACCTCAAA	1800
Db	2074	----CTTGTCTGTGCGCCCACTGATTTGTTCCGCAAGACATCCGGAAGCACATCTCCG	2129
OY	1801	TGCGGCTCTGGGCGCTGGCTACCGCCAAAGTGCATGACATACCCCTACAGGCTCTGG	1860
Db	2130	TGCGGCTCCGGTCCCTGGATTACACCCAGGTGACATGGTCGACTCCCGTATAGGCTTTGG	2189
OY	1861	CATTACCCCTGCACAGTTAATCTATACCATCTTCCAAAATAAGAGATATGTGGAGGGTT	1920
Db	2190	CATTATCTTGTACCATTAATTTACACCAATTTCAAGATGTAATGTGGAGGGGTC	2249
OY	1921	GAGCAGAGGCTCACGGCTGCATGTGCAATTTCACTCGTGGGAAATCGTTGCAATTTGAGAGAC	1980
Db	2250	GAGCAGAGGCTGGAAGCGGCGCTGCACATCGAGCGGGGCGAAACGCTGTGATCTGGAAGAC	2309
OY	1991	AGAAGCAGAAAGTCAATGTCTCTTTTGTGACCTTCACACAGAGATGGGCAATTTTAACT	2040
Db	2310	AGGACAGAGTCCGAGCTCAGCCCAATTTGTCTGTCTCACACACAGTGGCAGTCTTCG	2369
OY	2041	TGCTCTTACTGGGACCTGGCCGCTTGTGTGACATGTCTTCTTCACCTTCACCAAAATCT	2100
Db	2370	TGTTCTTTCACGAGCCCTGCGACGCTTGTGCCACGCGCTCATTCACCTTCACMAAAATTT	2429
OY	2101	GTGACGTAACATTCATGTATGGGCTTACACTGCGCCCTCAAAAATACATGTGCGATGG	2160
Db	2430	GTGACGTAAGTACTTGTACGGGTAGAGTTCMAAGATCGGCTCTGGGCCATTAATGTG	2489
OY	2161	GAGTGGGTAACTCTTATTTCTGTCTCTTACCGGACGCGCAGGATTTGGCGCTGCTATGG	2220
Db	2490	GAGTACGTCGTTCTCTGTCTCTCTCTGCTTCCAGACGCGCGGTCTGTCTGTGTGG	2549
OY	2221	ATGTCTACTTGTGTGGGCGAGCGGAGACAGACCTAGAGAGCTGTGATCTTGACAGCT	2280
Db	2550	ATGATGTTACTCATATCCCAAGCGAGCGGCTTTGAGAACTCAGTAATCTCAATGCA	2609
OY	2281	GCGAGCGGAGCTAGCTGCAATGGGCTCTCTAATTTTGTGCATCTTTTGTGGGCGCTGG	2340
Db	2610	GCAATCCCTGGCGGGAGCGACGAGCTTGTGTCTCTCTCGTGTCTTCTGCTTTGGCGTGG	2669
OY	2341	TACATCAAGGTCGGGTAGTCCCTTATGACTAATTTCCCTACCTGACGTGTGCTCTTT	2400
Db	2670	TATCTGAAGGGTATAGTGGGTGCCCGGAGCGGCTTACGCTTCTTACGGGATGTGGCTCTC	2729
OY	2401	AGCTTACTGTCTCTAGCAATTTGCCCCAACAGGCTTATGTCTTATGACGATCTGTGCAATGGC	2460
Db	2730	CTCTGCTCTCTGCGCGCTTGTCTCTCCTCAGCGGGCATACGCACTGACACAGAGGTGGCGCG	2789
OY	2461	CAGATAGGAGGGGCTCTGCTGATATATATCACTCTTTACTCTCACCCCCGGTATTAAG	2520
Db	2790	TCTGTGTGGGCGGCTTGTCTTGTGGGGTTAATGGCGCTGACTCTGTGTGCCATATTTACAG	2849
OY	2521	ACCTTCTCAGCCGGTTTTTGTGTGATGTGTCTATCTTTCAGACCTGGGGGAAGCTATG	2580

Db	2850	CGCTAACAACACTGGTGCATGTGGTGGCTTCAGATATTTCTGCACAGAGTAAGAACGCCAA	2909
Qy	2581	GTCCAGAGATGGGACCAACCTATCAGAGGTGGCGGTGGCCGTGATGGCATCATATGGGCC	2640
Db	2910	CTGCACGATGAGGTTCCTCCCTCCATACGTCCTGGGGGGGGCCGATGCCGTCATTTTACTC	2969
Qy	2641	GTCCGCATATTTCTACCCAGGTGTGTGTTTGACATTAACCAAGTGGCTTTGGCCGTCTT	2700
Db	2970	ATGTGTGTGTACACCCGACCTCGGTATTTTGATCATCAACAACTACTCTGGCATCTTC	3029
Qy	2701	GGGCTGCTTACCTTCCTAAAGGTCTTTGACGGCGCTGCCGTACTTTCGTACGGCTCAC	2760
Db	3030	GGACCCCTTTGGGATTTCTTCAAGCCACGATTTGGTTAAATGCCCTTACTTCGTGGCTTCAA	3089
Qy	2761	GCTCTACTGAGGATGTGCACCATGAGCAAGGACTCTCCGGGGGGCAGATACGTCCAGATG	2820
Db	3090	GGCCTTCTCCGGAATCTGCGCTACGCGCGAAGATAGCCGAGGTCACTTACGTTCAAATG	3149
Qy	2821	GCGCTACTAGCCCTTGGCAGGTGACTGGCACTTACATCTATGACCACTTCAACCCCTATG	2880
Db	3150	GCATCATCAAGATTAGGGGGCGCTTACTGGACCTATGTGTATTAACATCTCACCCCTCTT	3209
Qy	2881	TCCGATTTGGCTGCTGATGGCTTCGGGACCTGGCGGTCCGCCGTTGACCTATCATCTTC	2940
Db	3210	CGAGCTGTGGGCGCAACAACGGCTCGAGATCTGCGCGGTGGGTGTGGAACAAGTGTCTTC	3269
Qy	2941	AGTCCGATGAGGAAGAAAGTATGTCTGGGGGACCGGACAGCTGTCTTGTGGGGAATTT	3000
Db	3270	TCCGGAATGGAGACCAAGCTCATACGATGGGGGGCAATACCGCGGTGGGTGATCATC	3329
Qy	3001	TTTACACGACCTCCCGGTCCGCCCGCATTTGTGTGGGAGGTCTCTTGGCCCAAGCTGAT	3060
Db	3330	ATCAACGGCTTGGCCGTCTCGCCCGTATGGGGCCAGAGATATCTGCTTGTGGCCAGCCGAC	3389
Qy	3061	GGCTATACCTTCCAAAGGGTGAAGTCTTCTGGCCCCATCACTGCTTACGCCACAGACA	3120
Db	3360	GGAATGTGCTCCAAAGGGTGAAGGTCTTGCGCCCATCAACGGGTACGCCACAGACAGC	3449
Qy	3121	CGTGGCCCTTTTGGGACCATATGTGTGTAGATGACGGGGGGCCACAAGACAAACAGCT	3180
Db	3450	AGAGGCTCTCTAGGAGGTGTATATATACCAAGCTCTGACCGGGACAAAAACCAAGTGAAG	3509
Qy	3181	GGGGAATTTCAAGTCTCTGTCCAAGTCACTCAGTCTCTCCGGGAATCATCATCTCGGGG	3240
Db	3510	GTGAGGTCCAGATCTGTCTCAACTGTCTAACCCAACTTCTTGGGACGTGATCAATGGG	3569
Qy	3241	GTTTTGTGAGCTGTCTACCATGAGCTGGCAACAAGCTCTGGCCGGCTCACGGGGTCCG	3300
Db	3570	GTATGTGTGAGCTGTCTACACAGGGGGCCGGAACGAGACCATCGCATCACCAAGGGTCTT	3629
Qy	3301	GTCAACGAGATGTACTCAAGTGTCTGAGAGGGGACCTTATAGGTGGGCCACCCCTCTGG	3360
Db	3630	GTCAATCCAGATGTATACCAATGTGGAACAAGACTTGTGTGAGCTGGCCCGCTCTCAAGGT	3689
Qy	3361	ACTTAATCTTTGAGAGCCGTGACGTGTGTGAGACGGTGCACCTGTATCACTGTCAACGGGAAC	3420
Db	3660	TCCCGCTCATTTGAACCTCTGACCTTGCGGCTCTTCGGACCTTTTACTGTGTCAAGAGGCAC	3749
Qy	3421	GCTGATGTCAATCCCGGCTCGAAGACGGGGGACAAACGGGAGACGCTACTTCCCGGAGA	3480
Db	3750	GCCGATGTCAATTCCTCGTGCCTGGCGGAGAGTGAATGACAGGGGTAGCTGTTTGGCCCGG	3809
Qy	3481	CCTCTTTTCCACTTTGAAGGGGTCTCTACAGAGGCCCGGTGTATATGCCCCAGGGCCACGCT	3540
Db	3810	CCCATTTTCTTACTTTGAAGGCTCTCTCGGGGGGTCTCCGTGTGTGTGCCCCCGGGACAGGCC	3869
Qy	3541	GTCCGAGCTTTCCGGGACGCTGTGTGTCTCGGGGCGTGGCTTAAGTGCATATAGATTTTCATC	3600
Db	3870	GTGGGCTATTTCAAGGCGCGGTGTGTGACCCGTGTGATGGCTTAAGCGGTGTGACTTTATC	3929
Qy	3601	CCCGTTGAGACACTCGACATCTGTACCGGGTCCCCCACTTTATGTGACAAACAGACACCA	3660

Db	3930	CCTGTGGAGAACCTTAGAGACAACATGAGATCTCCCGGTTCACGAGCAACTCTCTTCCA	3989
Qy	3661	CCTGTGTGCCCCAGACCTTATCAGTGGAGTACTTGACATGCCCGGACTGGAGTGGAAAG	3720
Db	3990	CCAGACAGTGCCTCAGAGCTTCCAGGTGGCCACCTGCATGCTCCACCGGACGGAGCGTAAAG	4049
Qy	3721	AGACCAAAAGTTCTGTGGCATATGCTGCTCAGGGGTATAAGTCTAGTGGCTTAATCCC	3780
Db	4050	AGACCAAAAGTCTCCGCTGCGCATGCCAACCCAGGGCTACAAAGTGTGGTCTCAACCCC	4109
Qy	3781	TCAGTGTGCTGCACCTGGGGGTTTTGGGGCGTAACTGTCTAAGGGCAATGAGCATCAATCCC	3840
Db	4110	TCGTGTGCTGCAACCGCTGGGCTTTTGGTCTTACATGTCCAAAGGCCCAATGGGGTTGATCTC	4169
Qy	3841	AACATTTAGAGACTGAGTACGAGCTGTGACGACCGGGGCGCCCATCACTGACTCAATAT	3900
Db	4170	AATATCAAGACCGGGGTGAGAAACAATTACCACTGGCAGACCCCATCACTGACTCACTAC	4229
Qy	3901	GGCAAAATTCCTGCGCGCATGGGGGGCTGTGCGGGCGGCGCTACAGACATTCATATGTGAT	3960
Db	4230	GGCAAGTTCTCTTGGCGACGGCGGGGCTCAGAGAGTGTCTTAGACATATATTAATTTGGTAC	4289
Qy	3961	GAATAGCCATGCGCGTGACTTTACACCAATCCCTTGGCATTCGGAACAGTCTTGATCAAGCA	4020
Db	4290	GAGTGCCTCCACGGAATGCCACATTCATCTTTGGGATTCGGCACTGTCTTTGACCAAGCA	4349
Qy	4021	GAGACAGTGGGGTGCAGACTAATCTGTGTGGCTACAGCTACGCCCTCGAGTACATGACA	4080
Db	4350	GAGACTGGGGGGGCGAGACTGGTTGTGCTGCGCACTGTGTAACCTCCGGGCTCGTCACT	4409
Qy	4081	ACCCCTCCACCCCAATATAGAGAGGTGGCTTTGGGACAGAGGGCGAGATCCCCTTCTAT	4140
Db	4410	GTGTCCCACTCCATATCAGAGAGGTTCTCTGTTCACACCGGAGAGATCCCTTTTAC	4469
Qy	4141	GGGAGGGGGGATTTCCCTGTCTTACATCAAAAGGAGGAACAATTCGATTTCTGCACTTCA	4200
Db	4470	GGCAAGGCTATCCCTCTGAGGTGATCAAGGGGGGAGAACATCTATCTTCTGCACTCA	4529
Qy	4201	AAGAAAAGTGTGACGAGCTCGCGCGGCGCTTCGGGGTATGGGCTTTGAACTCAGTGGCA	4260
Db	4530	AAGAAAGAGTGGACGAGGTGCGCGCGGAGAGCTGTGCGATTGGGGATCAATAGCGGTGGCC	4589
Qy	4261	TACTACAGAGGTTGAGCGTCTCCGTAATACCACTCAGGGAGAGTGTGTGTCGCC	4320
Db	4590	TACTACCGGGCTTTGACGTGTCTGTCAATCCGACCAAGCGGAGTGTGTGTGTGTGCG	4649
Qy	4321	ACCGAGCGCTTCATGACAGGGGTATATCTGGGGACTTTGACTCGGATGAGCTGCAAGCTA	4380
Db	4650	ACCGAGTCTCTATATACGTGGCTTACCGGCACTTGCATCTGTATAGACTGCAACACG	4709
Qy	4381	GGGCTCACTCAAGTGTAGACTCTCAGTTTGAAGCCCAATTCACCATTAACACACAGATT	4440
Db	4710	TGTGTACCTCAGACAGTGGATTTCAGGCTTGACCTTACTTTACATTGAGAACACACG	4769
Qy	4441	GTCCCTCAAGACGCTGTCTCAGCTAGCCAGCGCGGGGTTCGACCGGTTAGGGGAAAGCTG	4500
Db	4770	CTCCCCAGAGAGTGTCTCTCAGACTCAACCGCGGGGCAAGGACTGCGAGGGGAAAGCCA	4829
Qy	4501	GGCATTTATAGGTATGTGTTCCATGCTGTAGCGAGCTCAAGAAATGTTTGAAGTGTAGTG	4560
Db	4830	GGCATCTTACAGATTTGTGGCACCGGGGAGGCGCCCTTCGGATGTTTGAATCTGCTGCTC	4889
Qy	4561	CTCTGTAGTGTACGACGACGAGGGCGCGATGTGTAGCTTACACATTCGAGACGACCC	4620
Db	4890	CTCTGTGAGTGTCTTACGACGGGGCTGTGTGTGTATGAGCTTACGCCCGCGGAGACTTCA	4949
Qy	4621	GTCAAGCTCAGGGCGGTATTTCAACGCGCGGTTTGCTGTGTGCAAGACCATTTTGAAG	4680
Db	4950	GTTAGGCTTACGAGCGTACATGAAACCCCGGGGCTTCCGATGTGCAAGACATTTTGA	5009
Qy	4681	TTTTGGAGGACGATTTCACCGGCTCAGACACATATGATGCCCATTTCTTTCCCAACA	4740
Db	5010	TTTTGGAGGGGGGTCTTTAAGGGCTCTACATATATATATATGATGCCCATTTCTTATCCAGCA	5069

QY	4741	AAAGCAATCGGGGGAAAAATTTCGCACTACTTAACAGCCTTACAAGGCTAAAGTGTGGCTAAG	4800
Db	5070	AAGAGAGGTGGGGAACTTTCCTTACCTGGTACCGTAACAAAGCAACGTGTGGCTAAG	5129
QY	4801	GCCAAAGCCCCCCCCCGTCTCTGGGAGGTCATGTGTGGAAGTGTGTGACTGCATCTCAAGCCC	4860
Db	5190	ACCTTCATGGGGCCAAACCCCTGCTATACAGATCGAGGCGCTGTTCAGAAATGAAGTACC	5249
QY	4921	CTCAACATCCCGTGAAGAAATACATGCGCCACCTGCATGCAAGCGCACTTGAAGTCATG	4980
Db	5250	CTGACGACCCCAATACCAATATATATATGATATGATATGATGCGCCGACCTGAGAGTGTCTC	5309
QY	4981	ACCAGCATTGGGCTTTGGCAGGGGAGTCTTGGCGGCGCTCGCCGCTATTTGCTTGGCG	5040
Db	5310	ACGAGACCTGGGTGCTGTGGGGGAGTCTCGGCTCTCTCGGCGCGGTATTTGCTGTCA	5369
QY	5041	ACCGGTGTGTTTGCATCATCTGGCGCGCTTGCATATTAACAGCGAGCGGTCTTGGCGCG	5100
Db	5370	ACAGGCTGCGGTGTATATGTGGGAGATTTGTTTGTCCGGGAAGCCGGCAATTATCTT	5429
QY	5101	GACAAAGAGTCTCTATAGAGCTTTTGTATGATGAGATGAGAAATGTGCTCTAGGCGGCT	5160
Db	5430	GACAGGAGGTTCTCTACAGAGATTTGATGATGAGATGAGATGTCTCTACAGACTTACCG	5489
QY	5161	CTCATTTGAAGGGGGACCGGATATGCGGAGATGCTGAATTCGAATTCGAAGCTTATTTG	5220
Db	5490	TACATTCGAGCAAGGGATGATGTCTGCTGAGAGTTCAAGAGAAAGGCGCTTCGGCTCTG	5549
QY	5221	CAGCAAGCTTCCAAACAGCTCAAGACATACAAACCACTGTGCAGGCTTATGAGCCCAAG	5280
Db	5550	CAGACCGGCTCCCGCCAGCAGAGATTATACCCCTGCTGTCCAGACCAACTGGCAGAAA	5609
QY	5281	GTAGAACAATTTCTGGGCCCAACACATGTGGAACTTCAATTAGCGGCATCCAAATCTTGCA	5340
Db	5610	CTCAGAGCTTCTGGGCGAAGACATGTGGAAATTTCAATGATGGAATACATATCTTGGCG	5669
QY	5341	GGACTATCAACATCGCCAGGGAGACCCCTGAGTACCTTCATATGAGGGTTCAATGTCCGCC	5400
Db	5670	GAGCTGTCAACGCTGCGCTGTGTACCCGCAATCTTCAATGATGAGCTTTTACAGCTGGCC	5729
QY	5401	CTCACAGTCCGCTGTCAACAGAGCACATATCTTCTCAACATTTTGGGGGCTGTGCTA	5460
Db	5730	GTCAACAGCCCACTAACCATCTGGCCAAACCTCTCTTCAACATATTTGGGGGGTGGGTG	5789
QY	5461	GCATCCCAATTGCACCAACCCCGGGGGGCCACTGGCTTTCGTTGTCAATGTGCTTATGAGGA	5520
Db	5790	GCTGCCCAAGCTCGCGCCCCCGGCGCGCTACCGCCTTTGTGGCGGCTTATGCTTACGTGC	5849
QY	5521	GCTGCCGAGGACAGTATAGCTTTAGGTAAAGGTGTAGTGAATCTTGGCAGAGGTATAGT	5580
Db	5850	GCGCGCATCGGACAGCTTGAACATGGGGAAGTCTCTGTGAATCTTTCGACAGGTATATGCG	5909
QY	5581	GCGGCAATTTGGGGGGCTCTGCTGCATTCAGATCATGTCTGGCGAGAAAGCCTTCATG	5640
Db	5910	GCGGGCGTGGGGGAGCTTTGTAGCAATTCAGATCATGAGCGGTGAGGTCTCCCTCCACG	5969
QY	5641	GAGCATGTGTCACATTGCTGCGCTGGGAATTCGTGTCGCGGTGTGCTTGTATGATGAGGAGTC	5700
Db	5970	GAGGACCTGTGTCAATCTGTCTGCGCCGCAATCTCTGCGCTGGAGCCCTTGTATGTGAGTGTG	6029
QY	5701	ATCTGCGGGGCATTTCTGCGCGCACAGCTGGGACCGGGGGAGAGCGCGCTTCCAAATGATG	5760
Db	6030	GTTCTGCGAGCAATATCTGCGCGGCAAGCTTGTGGCCGCGAGGGGGCAGTGCATATGAGATG	6089
QY	5761	AATAGACTCATTTGCTTTGCTTTCAGAGAGAAATACGTCGCGCCCAACCACTTACGTAGC	5820
Db	6090	AACGGCTAAATAGCTTCTGCTCCCGGGGGAAACCATGTCTTCCCAACGCACTTACGTGCG	6149

OY	5821	GAGTGGAGTGGGTGCGACGGTGTGACCACTACTTGGCTCCCTAACATAACAGACGCTG	5888
Db	6150	GAGACGAGTGCAGCGCCCGGCTCACTGCATCTACACAGCTACTACTGTAAACCGAGCTC	6209
OY	5881	CTCAGAAACCTCCACAACTGTATTCTAGGAGCTGCCCATTCCTATCGCGGGGCTGTGG	5948
Db	6210	CTGAGGCGCACTGTGCATCAGTGGATTAGCTCCGAGTGTACACTCCATGCTCCGGTTCCTGG	6265
OY	5941	CTCCGCGAGTGTGTGGGACTGTGGTTTGACCAATCTTAACAGACTTTAAAAATTTGGCTGAC	6000
Db	6270	CTAAGGAGCATCTGTGGGACTGTGATTAGCAGGGGTGTGACGACTTTAAGACTGCGCTGAAA	6329
OY	6001	TCCAAATTATTCGCCAAGATGCGCGGCTCCCTCTTTGTCTCTGTGCATAAAGGGGTACAG	6066
OY	6061	GCGGTGTGGGCGCGGCACTGGCATCTATGACCAACGCGTGTCTTGGCGGGCCAAATATCTCT	6120
Db	6390	GGGGTCTGGCCAGAGAGAGCGGCAATTATGCACACTGCTGTCCACTGTGTGAGCTGTGATCTCT	6449
OY	6121	GGCAATGTTCGCTTGGGCTCCATGAGAAATCACGGGGCTTAAGACCTTCATGAATATCTGG	6180
Db	6450	GGACATGTCAAAAAGGAGCGATGAGATCGTCGGTCTCTTAAGACCTTGAGGAACATGTGG	6509
OY	6181	CAGGGGACCTTTCTTAATATTGTTTACACGGAGGGCAGTGCCTGCCAAACCCGCGCA	6240
Db	6510	AGTGGAGAGTTTCCCATTTAAGCCCTACACACGGGCGCCCTGTACTCCCTTCTTGCGCGG	6569
OY	6241	AACCTTAAAGGTGGCATCTGAGAGGGTGGCGGCTCAGATACGGCGAGGTACCCAGAC	6300
Db	6570	AACCTTAAGTTTCGCGCTGGAGGGTGTCTGCAGAGAAATCGTGAATTAAGGGGGTGG	6629
OY	6301	GGGTCTATCCACTACATTAACAGAGACTCACTGATTAATCTTGAAGTCCCTGCAACTA	6360
Db	6630	GGGGACTTCCACTACATAGTACGGGTATGTACTGTCAATCTTTAAATGCGCTGCAGATC	6689
OY	6361	CCCTTCTCCCGAGTTCTTTTCTGGGTGAGCGAGTGCAGATCATAGTTTGGCCCCACA	6420
Db	6690	CCATTCGCCCGCAATTTTTCACAGAAATTTGACGGGGGCGCCCTACATATAGTTTGGCCCCCT	6749
OY	6421	CCGAAGCGGTTTTTCCGGGATGAGGTCGTCTGTCGGTGGGCTTAATTCATTTGTGTCTC	6480
Db	6750	TGCAGAGCCTTGTCTGCGGGAGAGGTATATTCAGAGTGAATCTCACAGATACCCGGTG	6809
OY	6481	GGGTTCGCAAGTCTCTTGTGCGACCCCTGAACCCGACACAGACGTAATGATTCATGCTAA	6540
Db	6810	GGGTTCGCAATTAATCTTGTGGAACCCGAAACCGAACGTAGCCGTGTGATCGTCAATGCTCACT	6869
OY	6541	GATCCATCTCATTCACGGCGGAGACTGACAGCGGGCGGCTTTAGCGCGGGGTCAACCCCA	6600
Db	6870	GATCCCTCCCATATTAACAGACAGAGCGGCGCGGAGAAAGGTTGGGAGAGGGTCAACCCCT	6929
OY	6601	TCCAGAGCAAGCTCTCTGCGCGACGCAAGCTATTCGGACCATGCTGTGCGACCACTGCACC	6660
Db	6930	TCTATGGCAGACTCTCTGCGCGACGCAAGCTGTGTCGCTCATCTCTCAAGGCAATCTTGACAC	6989
OY	6661	ACCACGCGCAAGCCATGATGTGTGACATGTGGATGTCTAACCT-----GTTC	6708
Db	6990	GCCAAACATGACTCCCTGTACGCGGAGCTTCAAGAGGCTTAACCTCTGTGAGGCGAGAG	7049
OY	6709	ATGGGGGGCGATGTGACTCGGATAGAGTCTGGGTTCAAAGTGTGTCTTGTGACTCTCTC	6768
Db	7050	ATGGGGCGGCAATCAACAGGGTTGAGTCAAGAAACAAAGTGGGATTTCTGTGACTCTCTC	7109
OY	6769	GACCCAAATGTGTGGAAGAAAGAGCACTTGAAGCTTGTGATACATCAAGATACATGCTC	6828
Db	7110	GATCCGCTTGTGGCAGAGAGATGATGAGGGAGGCTTCCGTAACCGCAAAATTTCTGGCG	7169
OY	6829	CCCAAGAAAGTGTCCCAACGACTTTACCGGCTGTGGGCAAGGCTGTATTAACACCAACG	6888
Db	7170	AAGTGTGGAATTTGCCCCGGGCGCTGCGCTTTTGGGGCGGCGGCACTAACACCCCGG	7229
OY	6889	CTTGTGATCTGTGAAAAAGGCTCAGATTACCAACCGGCACTGTGGCGGCTGTGCTCTC	6948

Db	7230	CTAGTAGACAGTGGAAAAAGCTGACTACGAAACCACTGTGGTTCATGGCTGCCCCGTA	72893
Qy	6949	CCTCCTCCCTAGAAAAAACCACGACCTTCCCAAGAGACCCGACAGTGGGCTTAAGT	70089
Db	7230	CCACTCCACGGTCCCTCTCTGTGCTCCGCTCGGAAAAAGCGTAGGTGGTCTTCAC	73458
Qy	7009	GAGCATCTCATAGAGATGCCCTTTCACAGCTGGCTGAATTAAGTCTTTGGCCAGCCCC	70666
Db	7350	GAATCAACCCCTATCTACTGCTTGGCCAGAGCTTGCACAAAGATTTTGGCAGCTCTCA	74059
Qy	7059	CCAAGCGGCGATTCAAGGCTTTCCACGGGGGCGGGCGGTGCAGATTCCGGCAGTCAAGC	71222
Db	7410	ACTTCCGCAATTACGGGCAATACGCAACATCTCTGAGCCCGCC-----	74575
Qy	7129	CCTCCGTGATGATTGGACCTTTTCGGAGACAGGTTCCATCTTTCATATCCCCCTCGAG	71889
Db	7458	CCTTCTGGCTCCCCCCCGACCTCCGACGCTTAAGTCTATTCTTTCATATCCCCCTCGAG	75171
Qy	7189	GGGAGCTTGGAGATCGACGCTTGAGCCTGAGCAGTGAAGCCCCCCCCCAG	72488
Db	7518	GGGGAGCTGGGGATCCGGA-----	75375
Qy	7249	GGGGGGGTGGCAGCTCCCGCTCGGACTCGGGGTCTGTGTACTTGTCTCGAGAGGAC	73089
Db	7538	-----TCTCAGCGACGGGCTCATGTGTGACGGTCAAGTAGTGGGGCGACACG	75833
Qy	7309	GACTCCGTCGATGTGCTGCATGTCATATCTCGGACGGGGGCTTAATTAACCTCTGT	73689
Db	7584	GAAATGTCTGTGTGCTCTCATGTCTTATTCTCGAGACGGCGACTGTCAACCCCGTGC	76433
Qy	7369	AGTCCCGAAGAGAGAAAGTTACCGATTAACCCCTTGAGCACTCCCTGTTCGATATCAC	74289
Db	7644	GCTGGGAAAGAAACAAAACCTGCCATCAAGCACTGAGCAACCTGTGTCTACGCCATCAC	77039
Qy	7429	AACAGAGTGTACTGTACCACAAAGAGCCCTCACTAAGGGCTAAAAAGTAACTTTT	74889
Db	7704	AATCTGGGTATTCACACACTTTCACGCACTGTTCGCAAAAGCAGAAAGTCACAATT	77633
Qy	7489	GATAGGATGCAAGTGTGACTCTCTTACTAGACTCAGCTTAAAGGACATTAAAGCTAGCG	75489
Db	7764	GACGACTGCAAGTTCGTGGACAGCATTTACCAAGACGTGCTCAAGAGGCTTAAGCAGCG	78229
Qy	7549	GCTTCGAAGTACACCGCAAGGCTCCTCAACATGAGAGAGGCTTGCAGTTAACCCACC	76089
Db	7824	GCGTCAAAAGTGAAGGCTTAATTCTCTATCCGTAAAGGAAGTTGACGCTACGCCCCCA	78833
Qy	7609	CATTCTGCAAGATCTAAATATATGGGTTTGGGGCTAAGAGGTCCGACGTTGTCCGGAGG	76689
Db	7884	CATTGAGCAATCCAAAGTTGGCTATGGGGCAAAAGACGTCCTGGCCATGCAAGAAAG	79433
Qy	7669	GCCGTAAACCAATCAAGTCCGTGTGGAGAGACCTCTCGAGAGACTCAGAAACACCAATT	77289
Db	7944	GCCGTAGCCCAATCAACTCTCGTGTGAAAGAACTTTCTGGAAGACAGTGAACACATA	80033
Qy	7729	CCCAACAACATTATGAGCCAAAAATGAGGTGTTCTGCGTGAACCCACCAAGGGGGGCAAG	77889
Db	8004	GACACTACATCATAGCCCAAGAGAGAGAGTTCGTGCTTCAAGCTTGAGAGAGGGGGTGT	80633
Qy	7789	AAAGCAGCTGCTTATGTTTAACTTGACCTTCGGCGTCAAGGTCGTGAGAGATGGCC	78489
Db	8064	AAGCAGCTGCTCATGTGTGTTCCCGACCTTGGCGTGGCGTGTGGAGAAATGGCC	81233
Qy	7849	CTTATATGACATTACACAAAACTTCTCTCAGCGGTGATAGGGGCTTCTTATGATTTCCAG	79089
Db	8124	CTGTACACGCTGTGTAACGACCTCCCTCGCGTGTGATGGGAAGCTCTTACGATTTCCAA	81833
Qy	7909	TATTCCTCCGCTCAGCGGTGAGATTTTCTCTTGAACATAGCGCGGAAAAAGAGACCTT	79689
Db	8184	TACTACACAGGACAGCGGGTTGAATTCCTGTGTGAACGCTGGAAGTCCAGAAACACCCG	82433
Qy	7969	ATGGGTTTTTGTATGATACCCGATGCTTTGACTCAACCGTCACTGAGAGACATCAGG	80289

Db 8244 ATGGGGTTCTGTATGATACCCGCTGTTTGTGACTCCAGTCACTGAGAGCGAATCCGT 8303
QY 8029 ACTGAGAGTCCATATATCGGGCTGCTCTTGGCCGAGGAGGCCACACTGCAATAC 8088
Db 8304 ACCGAGAGGCAATTTACCAATGTTGTGACTGGAACCCCAAGCCCGCTGGCCATCAAG 8363
QY 8089 TCGTAACTGAGAGACTTTTACGTGGAGGGCTTATGTTCAACAGCAAGGCCCAACTGC 8148
Db 8364 TCCCTCACTGAGAGGCTTTATGTTGGGGGCGCTCTTCAATTCAAAGGGGGAATAATGC 8423
QY 8149 GGGTACAGGGGTTGCGCGCCGAGGGGGTGTCAACCTGAGCAATGGGGAACACATACA 8208
Db 8424 GGGTACAGAGTCCCGCGAGCGGCTACTGCAACTAGCTGTGTAAACCCCTACT 8483
QY 8209 TGCTACGTGAAGGCTTAGCGGCTTGTAAAGCTCAGAGGTAATCGCGCCCAATGCTG 8268
Db 8484 TGCTACATCAAGGCCCGGAGGCTGTGAGCCGCAAGGCTCCAGAGCTGCAACATGCTC 8543
QY 8269 GTATGCGGCGATGACTTGTGTGATCTCAGAAAAGCAGGGGACGAGAGAGAGCGG 8328
Db 8544 GTGTGCGCGACGACTTATGTTATCTGTGAAAGTCCGGGGGTCCAGAGAGAGCGGCG 8603
QY 8329 AACCTGAGAGCTTCAAGAGGCTATGACCAAGTATCTGCTCTCAATGTGTGAGCGCTG 8388
Db 8604 AGCTGAGAGCTTCAAGAGGCTATGACCAAGTATCTGCTCTCAATGTGTGAGCGCTG 8663
QY 8389 AGACCGAGTATGATCTGAGCTGATTAACATCTGCTCTCAATGTGTGAGCGCTG 8448
Db 8664 CAACCAATATCGACTTGAAGCTTATTAATCATGCTCTCTCAAGTGTGAGCGCTG 8723
QY 8449 GGGCCACAGGCGCGCGAGATCTACCTGACGAGAGACCTTACACTCAATGCGCGG 8508
Db 8724 GACGCGCTGGAAGAGGGGTCTACTACTTACCTGACCTTCAACACCCCTGCGCAGA 8783
QY 8509 GCTGCTGAGGAAAGATTAGCACTCCCTGTGAATTCATGAGCTGGGAAATCATCATCAG 8568
Db 8784 GCGCGTGGAGAGAGAGACACACTCAAGTCAATTCCTGGCTAAGGCAATATATATG 8843
QY 8569 TACGCGCGGACCATATGAGGCTCGCATGCTGTATGACACACTTCTTCCATTTCTATG 8628
Db 8844 TTTGCCCCCAGCTGTGGGCGAAGATGATGATGATGATGATGATGATGATGATGATGATG 8903
QY 8629 GCTCAAGACAGCTGAGACAGAACTCACTTTGAGATGTAAGAGGCGGTACTGCGTG 8688
Db 8904 GCCAGGATGATGCTTGAACAGGCTCTTAAGTGAATCTAAGAGCTGCTACTCATTA 8963
QY 8689 AGTCCCTTGAACCTCCAGCTATTAATGAAAGTTAATGAGGCTTGAACGCTTTTCTGTC 8748
Db 8964 GAACCACTGATCTTACCTCAATCAATTCAAAGACTCAAGGCTCAAGCGCATTTTCACTC 9023
QY 8749 CACAATATCACTCCCAAGAACTGACAGGAGTGGCTTCAAGCCCTCAGAAAACTTGGAGCG 8808
Db 9024 CACAGTATCTTCAAGGTGAATCAATAGGAGTGGCGGATGCTCTCAAAAACTTGGAGCG 9083
QY 8809 CCACCCCTCAGAGGCTGAGAGAGCGCGGACAGTCAAGGAGGCTCCCTATCTCCGT 8868
Db 9084 CCGCCTTGGAGCTTGAAGACACCGGCGCGAGCGCTCGCGCTAAGCTTCTGTCCAGA 9143
QY 8869 GGGGGAGAGCGCGCTTGTGCGTGTGATATCTTCAATTTGGGCGGTGAAGAACCAACTC 8928
Db 9144 GGAAGCGAGGCTGCAATATGTGGCAAGTACTCTTCAACTGGGAGTGAAGAACCAACTC 9203
QY 8929 AAATCTCACTCATTTGCGGAAAGCGGCTCTGGAATTTATCAGCTGGTTTCAACCGTGGG 8988
Db 9204 AAATCTCACTCAATATAGCGCGCGCTGGCGGCTGGAATTTGTCCGGTTTCAAGGCTGGC 9263
QY 8989 GCGGCGGAGGCGCAATTTATCAAGCGGTGTGCGGTGCGGACCCCGCTTATTTGCTCTTT 9048
Db 9264 TACAGCGGAGGAGCAATTTATCAAGCGGTGTCTCATGCGCGCGCTGCTGTGCTT 9323
QY 9049 GGGCTATCTCTTCTTTTGTAGGGTGAAGCTTTTCTCACTCCCGCTCG 9098
Db 9324 TGCCTACTCTGCTGCTGAGGGGTAGGATCTTACTCTCTCCCAACCG 9373

RESULT 11
US-09-014-416-2
; Sequence 2, Application US/09014416
; Patent No. 6153421
; GENERAL INFORMATION:
; APPLICANT: Yanagi, Masayuki
; APPLICANT: Bulh, Jens
; APPLICANT: Emerson, Susanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
; FILE REFERENCE: 20264276
; CURRENT APPLICATION NUMBER: US/09/014,416
; EARLIER FILING DATE: 1998-01-27
; EARLIER APPLICATION NUMBER: US 60/053,062
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 9599
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; US-09-014-416-2

Query Match 45.3%; Score 4121.6; DB 3; Length 9599;
Best Local Similarity 66.4%; Pred. No. 0;
Matches 6046; Conservative 0; Mismatches 2974; Indels 90; Gaps 5;

QY 1 ATGAGCAAAATCTTAACTTCAAGAAAAACCAAAAGAAACCAACCTGCGCCACAA 60
Db 342 ATGAGCAAAATCTTAACTTCAAGAAAAACCAAAAGAAACCAACCTGCGCCACAG 401
QY 61 GACGTTAATTTCCGGGCGCGCGCGAGATGCTGTGGAGTACTGTTGGCGGCAAG 120
Db 402 GACGTTAATTTCCGGGCGCGCGCGAGATGCTGTGGAGTACTGTTGGCGGCAAG 461
QY 121 GGGCCAGATTGGGTGTGCGCGAGCAAGAAAGACTTGGAGCGGTCCAGCCACTGGA 180
Db 462 GGGCCAGATTGGGTGTGCGCGAGCAAGAAAGACTTGGAGCGGTCCAGCCACTGGA 521
QY 181 AGGCGCAGCCCATCTTAAAGATGGCGCTCCACTGSCAAATCTTGGGAAAAACAGA 240
Db 522 AGAGCTAGCATATCCCAAGGCAAGTGGCGGAGGAGCACTGGGCTGAGCCGCGG 581
QY 241 TACCCCTGCGCCCTATACGGGAATGAGGACTGCTGGGCGAGAGTCTCTGTCGCC 300
Db 582 TACCCCTGCGCCCTTATAGCAATGAGGATGGGTGGGCGGAGTGGCTCTGTCGCC 641
QY 301 CGAGTTCGCGTCCCTCTTGGGCGCCCAATGACCCCGGCAATAGTGGCGCAAGTGGGT 360
Db 642 CGTGGCTTGGGCTTATAGTGGGCGCCCAAGAGCCCGGCGATAGTGGCGCAATTTGGGT 701
QY 361 AAGTATGATATCCCTTAAAGTGGGCTTGGCGCACTCATAGGAGTATCTCTGTCTG 420
Db 702 AAGTATGATATCCCTTAAAGTGGGCTTGGCGCACTCATAGGAGTATCTCTGTCTG 761
QY 421 GGGGCGCGTGGGCGGCGTGGCAAGCTTGGCGCAATGGGTGAGTCTTGGAGAGAG 480
Db 762 GGGGCGCGTGGGCGGCGTGGCAAGGCGTGGCGCAATGGGTGAGTCTTGGAGAGAG 821
QY 481 GGGGTAATTTTGAACAGGAACTTACCGGTTGCTCTTTCTATCTTCTTGTGAGCG 540
Db 822 GGGGTAATTTTGAACAGGAACTTACCGGTTGCTCTTTCTATCTTCTTGTGAGCG 881
QY 541 CTGCTGTCTGATCAACACCCGCTCTCGCTGCGAAAGTGAAGAAATCAATACCGCG 600
Db 882 CTGCTGTCTGATCAACACCCGCTCTCGCTGCGAAAGTGAAGAAATCAATACCGCG 941
QY 601 TATATGTATCAAGCACTGACCAATGACAGATTAAGTGGAGTCCAGAGCTGTGTC 660
Db 942 TACCATGTACCAATATTTGCTTAACTGAGATTTGTATGAGGCGGCGGATGCAATC 1001

OY	661	TTCCACGTCCTCCGGGTGCTCCCGGCGAGAAAGTGGGGAAATGCATCTCACTGCTGGATA	720
Db	1002	CTGCACACTCCGGGGGTGTGTCTTCCTTGCGCTTGCGACAGGGGTAAAGCGCTTCAGGTGTTGGGTG	1061
OY	721	CCGGCTTCACCGAAATGTGGCCGCTGACGGGCCCCGGGCCCTTCACGACAGGGCTTTCGGACG	780
Db	1062	GGGGTGAACCCCAAGGTGGCCACACAGGAGCGGCAAACTCCCAACAGCAAGCTTCGAGCT	1121
OY	781	CACATCGACATGTTGTGATGTCGGCCACGCTGTGCTCTGCTTCATGTTGGGGGACCTC	840
Db	1122	CATATCGATCTGTGTGCGGGAGGGCCACCTCTGCTGGGCCCTTCAGTGTGGGGGACCTG	1181
OY	841	TGGGGTGGGGGATGCTCGACAGCCCAATGTTCATGTGTGCGCGCGACACACTGCTTT	900
Db	1182	TGCGGGTCTGCTTCTTCTGTGTGGTCACTGTATTACTTCTCTCCAGGCGGCACCTGGACG	1241
OY	901	GTCGAAGACTGCAATTGCTCTCATCTTACCCCTGTGACATCACTGGACACCGCATGAGATG	960
Db	1242	ACGCAAGACTGCAATGTGTTCTATCTATCTCCGGCCATTAAGGGGTCACTCGATGGCATTG	1301
OY	961	GACATGATGATGAATGTGTGCGCCACGGCTACCATGATCTTGGCGTACGATGCGTGTCT	1020
Db	1302	GATATGATGAAGAACTGGTCTCCCTTACCGCACGTTGGGTGTAGCTCAAGCTGCTCCGATC	1361
OY	1021	CCCGAGGTCATTATAGACATGATTAAAGGGGGCTCATTTGGGGGCGTCAATTTGGCTTGGCC	1080
Db	1362	CCACAGCCATCATGAGACATATGCTGTGTCTCACTGGGGAGTCTGCGGGCAATATGCG	1421
OY	1081	TACTTCTCTATGACAGGAGCGGTGGCGGAAATGTCGTTGTCACTCTTCTGTGGCCGCGGG	1140
Db	1422	TATTTCTCATGATGGGGGAATGGGCGAAGGTCCTGTGATGTGCTGTCTATTTTGGCGGG	1481
OY	1141	GTGACCGGCGCACCCATATCTGTTGGGGGTTCTGCGCGGACATCCGGGCGCTTCAAC	1200
Db	1482	GTCACCGGGAACCCACGTCAACCGGGGGAAATGCGCGGCCACACGCGCTGGCTTGT	1541
OY	1201	AGCTTATTTGAATGAGGGGCCGACAGGAGAAATTCAGAGCTGTTAACACCAATGGAGAGCTG	1260
Db	1542	GGTCTCTTACACACAGGGCGCCAGAGAGAAATCCAACTGATCAACACCAAGGCAATTTGG	1601
OY	1261	CACATCAACCGCACCGCCCTGTAAGCTGCATGACTCTTTCGACACCGGCTTTATCGCGTCT	1320
Db	1602	CACATCAATACAGGGCGCTTGAAATTTGCAATGAAGCTTTAACACCGCTGTGTATGACAGG	1661
OY	1321	CTGTTCTTACACCCACAGCTTCAACTCGTCAAGATGTCCGAAAGCATGTCCGCTGCGC	1380
Db	1662	CTCTTCTATCAACAACAAATTCACACTTTCAGGCTGTCTCTGAGAGATTTGGCCAGCTGCGGA	1721
OY	1381	AGTATCGAGGCGCTTCCGGGTGGGAATGGGGGCGCTTTCGAAATATGAGAGTAATGTCAACAA	1440
Db	1722	GCGCTTACCAATTTTGGCCAGAGGCTGGGGGTCTTATCAAGTTATGCCAACGAGAGCGGCTC	1781
OY	1441	CCAGAGATATAGAACCTTATTTGTGTGCATCAACCAACAGCAGTGTGGCGTGTCTCC	1500
Db	1782	GACGA-----ACGGCCCTTACTGTGTGGACATACCTCCAAAGACCTTGTGGCATTTGTGCC	1835
OY	1501	GCGAAGACTGTGTGTGGGCCGAGTGAATCTGTTTCAACCCCAAGCCAGTGTGTATGTGGCAGC	1560
Db	1836	GCAAAAGACGTGTGTGGCCGCGGTATATTGCTTCACTCCCAAGCCCCGTGTGTGTGGGAAGC	1895
OY	1561	ACCGACAGGCTTGGAGCGCCACTTACACGTGGGGGAGATGAGACAGATGTCTTCTCTA	1620
Db	1896	ACCGACAGGTGGGGCGCGCTCACTACAGCTGGGGTGAATGATACGAGATGTCTTGTCTC	1955
OY	1621	TTGAACAGACTCGAACCAACCGCTGGGTGCATGTTTGGCTGACAGTGTGATGAATCTTCT	1680
Db	1956	CTTAAACAACACAGGCCCAACCGCTGGGCATATTGTTGGTTGTATCTGGATGAATCTCACT	2015
OY	1681	GGCTAACACCAAGACTTGGCGGGGCAACACCTGCGGTACTAGAGCTGAATTCAGCGCAGC	1740
Db	2016	GGATTCACCAAGTGTGGGAGCGCCCTTGTGTCACTCGAGAGGGGTGGGCAACACA--	2073

OY	1741	ACGACCGCTGTGATGCCCAAGCATGTGTTTAAAGAAAGATCTGATTAACATTTACCTCAAA	1800
Db	2074	----CTTGCTCTGCCCCCATGTATGCTTCCGCAACAATCCGAGCCCAATATCTTCGG	2129
OY	1801	TGCGGCTCTGGGCGCCTGAGCTCAAGCCAAAGTGCTGATGACTAACCCCTACAGGCTTGG	1860
OY	1861	CATTACCCCTGACACAGTTAACTATACATCTTCAAAATAAGATGTAATGTGGAGGGCTT	1920
Db	2190	CACATACCTTGATCCATCAATCAATACACCATATTCAAAATCAGAGATGTACGTGGAGGGGCT	2249
OY	1921	GAGACAGGGCTCAAGGCTGATGTGAATTTAATCTGTGGGGATGCTGTCAACTTGGAGAGAC	1980
Db	2250	GAGACAGGGCTGGAGGGCGCTGTCACTGACGCGGGGGCAAGCGCTGTGATCTGGAAGAC	2309
OY	1981	AGACACAGAAGTCAACTGTCTCTCTTTTGTGCACTCCACCAAGGAATGGGCCATTTTACT	2040
Db	2310	AGGACAGAGTCCGAGCTCAGCGCGCTTCTGTCTGTCCACACACAGTGGCAGGTCTTCG	2369
OY	2041	TGCTCTTAATCTGACCTGCGCGCTTGTGCACTGTCTTCTGCACTTCCACCAAAATATC	2100
Db	2370	TGTTCTTTCAAGACCTCCAGACCTTGTCCACCGGCTCTATCCACTCCACACAGAAATTT	2429
OY	2101	GTGACAGGTCAATTCATGTAATGCGCTATCACTGCGCTCAACAAATATCATGTCCGATG	2160
Db	2430	GTGACAGGTGAGTACTGTATCGGGGTAGGGTCAAGCATCGCTCTGGGCAATTAAATGG	2489
OY	2161	GAGTGGGTAATACTCTTATTCCTGCTCTTATAGCGGACGCCAGGGTTTGCCTGCTTATGG	2220
Db	2490	GAGTACGTCTGTCTCTCTGTTCTTCTGTCTGTGACAGCGCGCGCTCTCTGTGTTGG	2549
OY	2221	ATGCTCATCTTGTGTGGCCAGGCCGAAGACAGCATAGAGAAGCTGTCAATCTTGACGCT	2280
Db	2550	ATGATGTTAATCATATCCCAAGCGGAGCGGCTTTGGAGAACCTGTAATATCTCAATGCA	2609
OY	2281	GCGAGCGCAGTACTGCAATGCGTCTCTATATTTTGTATCTTTTTCGTGGCGCTTGG	2340
Db	2610	GCATCCCTGCGCGGGAGCGACCGGTCTTGTGTCTCTCTGTTCTTCTCTTGTGCGTGG	2669
OY	2341	TACATCAAGGCTGGGGTAAAGTCCCTTTAGCTAACATTTCCCTCATGCGCGTGTGCTTT	2400
Db	2670	TATCTGAAGGATAGTGGGTGGTCCGAGCGGCTATACGCTCTTACGGGAATGTGCTCTC	2729
OY	2401	AGCTTACTGCTCCAGCATTTGCCCAACAGGCTTATGCTTATGACGATCTGTGCATATGC	2460
Db	2730	CTCTGTGCTCTGTCTGGGCTTGTGCTCAGCGGGCATATGCACTGAGACAGAGATGGCCGG	2789
OY	2461	CAGATAGAGCGGCTGTCTGTGTATGATCACTCTTTTACTCATCCCCCGGGTATAG	2520
Db	2790	TCTGTGTGGCGGCTGTGTTCTTGTGTGGGTTATAGCGCTGACTGTGTGCATATTTACAG	2849
OY	2521	ACCTTCTCAAGCCGCTTTTGTGTGTGTTGTGCTATCTTCTGACCCCTGGGGAAAGTATG	2580
Db	2850	CGCTATATCAAGCTGTGTGCATGTGTGTCTTCAGATTTTCTGACAGAGTAGAAGCCCAA	2909
OY	2581	GTCCAGAGATGGGACCACTTATGACAGTCCGCGGTGGCGGTATGTGATATATGGGCC	2640
Db	2910	CTGACAGTGTGGGTTCCTCCCTCAACGTCCGGGGGGGGCGATGCCGATCTTATCTC	2969
OY	2641	GTGCGCATATTTTCAACCAAGTGTGTGTGATCAATTAACAATGTGCTCTTGGCGGTCTT	2700
Db	2970	ATGTGTGTAGTACACCCGACCCCTGTGTATTGTAGATCACTCAAACTACTCTGTGGCATCTTC	3029
OY	2701	GGGCTGTGTTACCTCTTAAAGATGCTTGAACGCGGTGCCGTAATCTTGTACAGGCTCAC	2760
Db	3030	GGAGCCCTTGTGATTTCTTCAAGCCAGTGTCTTAAAGTCCCTTACTTCTGTGCGCGTTCAA	3089
OY	2761	GCTTACTGAGATATGTSCAATATGGCAAGCATCTCGCGGGGGGAGTATGTCAGATG	2820
Db	3090	GCGCTTCTCCGATATCTGCGCTACGCGGGAATAGCCGGAAGTCAATTAAGTGCAAATG	3149
OY	2821	GCGCTACTAGCCCTTGGCAGGTGATCTGSCATTTATATGACCAACTCACCCCTTATG	2880

Db 3150 GCCATCAATCAAGTTAGGGGCGCTTAAGTGGACCTATGTGTATACATCTCAACCCCTCTT 3209
Oy 2881 TCGATTGGGCGTCAAGTGGCGTGGGACCTGGGCGTGGCGCTTGGAGCCATATATCTTC 2940
Db 3210 CGAGCTGGGCGACACAGCGCTGGAGTCTGGCGGTGGTGGAAACAGTGTCTTC 3269
Oy 2941 AGTCGATGAGAAAGATCATTTGTCTGGGAGACGAGACAGCTGTCTTGGGAGATT 3000
Db 3270 TCCGAAATGAGAACCAAGCTCATCGTGGGGGAGATACCGCGCGTGGGTGATCATC 3329
Oy 3001 TTACACGACCTTCCGCTGTCCGCGGACTTGTGGAGAGTCTCTCTTGGCCAGCTGAT 3060
Db 3330 ATCAACGGCTTGGCGCTCTGGCCGTAGGGGCGAGAGATATCTTGGGCGCAGCGAC 3389
Oy 3061 GGCCTATCCTCAAGGGGTGAGTCTTCTCGCCCATCATCTGTCTTACGCCACAGACA 3120
Db 3390 GGAATGTCTCCAAAGGGGTGAGTGTCTGGCGCCATCATCGGGTACGCCACAGACG 3449
Oy 3121 CGTGGCTTTTGGGACCATAGTGTGAGATGACGAGGCGCGACAAAGACAGAGCT 3180
Db 3450 AGAGGCTCTTAGGGTGTATATACACAGCTGACTGGCGGGAACAAAACCAAGTGGAG 3509
Oy 3181 GGGGAAATTCAGTCTCTGTCCACAGTCACTAGTCTTCTCTGGAAATCATCTCGGGG 3240
Db 3510 GGTGAGGTCCAGATGTGTCACTGTCAACCAACCTTCTCTGGCAACGTGCATCAATGG 3569
Oy 3241 GTTTGTGAGACTGTCTACCATGAGCTGGCAACAAAGCTTGGCGGCTCAAGGGTCCG 3300
Db 3570 GTATCTGAGCTGTCTTACACGAGGCGGAAACAGAGACATCGATCAACCAAGGTTCT 3629
Oy 3301 GTACGACAGATGTACTCTCAGTGTGAGGGGAGCTTATAGTGGGTGGCCGCCCTCTGG 3360
Db 3630 GTCATCCAGATGTATACCAATGTGSAACAAAGCTTGTGGCTGGCCGCTCTCAAGGT 3689
Oy 3361 ACTAAATCTTTGAGACCGTGTGACGTTGTGAGCGGTGCACTGTACTGTGTACCGCGAAC 3420
Db 3690 TCCCGCTATTGACACCTGTGACCTGCGGCTCTCGACCTTGTACGTGGTACAGAGGAC 3749
Oy 3421 GCTGATGTATCTCCGGCTCGAAGACGCGGAGCAAAACGGGAGCGCTACTCTCCCGAGA 3480
Db 3750 GCCGATGTATCTCCGTGGCGCGGAGAGTGTATAGACAGGGGTATGCTCTTCCGCCCG 3809
Oy 3481 CCTCTTTCACCTTGAAGGGGTCTCAAGAGGCGCGGTGTATATGCCAGGGGACACGCT 3540
Db 3810 CCCATTTCTCTACTTGAAGGCTCTCGGGGGGTCTGGTGTGGCCCGCGGACACGCC 3869
Oy 3541 GTCCGAGTCTTCCGGGACGTTGTGTCTGTGGGGCGTGGTAACTCATATGATTTTATC 3600
Db 3870 GTGGGCTATTCAAGGCGCGGTTGTGACCCGTGAGTGTAAAGCGTGGACTTTATC 3929
Oy 3601 CCCGTTGAGACACTCGACATCGTCAAGCGGTCTCCCGACCTTTATGACAAACAGACACA 3660
Db 3930 CCTGTGAGAGAACTTAGGACAAACATGAGATCTCCGGGTGTTCACGAGAACTCTCTCA 3989
Oy 3661 CCTGTGTGCCACAGACTTATCAGGTCTGAGAGTCTTTCAGCTTGAAGCAACAGACACA 3720
Db 3990 CCAGAGGTGCCCAAGAGCTTCCAGGTGGCCCACTGATGTCTCCACCGGACAGGTAAAG 4049
Oy 3721 AGCACCAAGTCTCTGTGCAATATGTCTCTCAGGGGTATTAAGTGTATGTCTTAATCCC 3780
Db 4050 AGCACCAAGTCTCCGCTGTGCTGACGACCCAGGGCTTACAGAGTGTGTGTCTCAACCCC 4109
Oy 3781 TCAGTGTGCTGCCACCTGGGGTTTGGGGGTATCTTGTAAAGGACATGGGATCAATCCC 3840
Db 4110 TCTGTGTGCTGAACGCTGGGCTTTGTGTCTTATCATATGCAAGGCCCAATGGGGTTGATCTT 4169
Oy 3841 AACATTAGACTGTAGTCAAGACTGTGACACCGGGGCGCCATCATCGTACTTCCACATAT 3900
Db 4170 AATATCAAGACCGGGGTGAGAAATTAACACTGGACGCCCATCACTTCCACCTTAC 4229
Oy 3901 GGCAAAATCTCTGCGGATGGGGCTGTGTGGCGGCGCTTACAGATCATATATGTGAT 3960

Db 4230 GGCAAATCTTGTCCGACGCGGGGTCTCAGAGGTGCTTATGACATAATTAATTTGTGAC 4289
Oy 3961 GAATGCGATCGCGTGGACTCTACCAACATCTTGGCATGGAAAGATCTTGTATCAAGCA 4020
Db 4290 GAGTGCATCTCCACGGAATGCATCATCTTGTGGGATGTGGCATGTCTTATCCAAAGCA 4349
Oy 4021 GAGACAGTGGGGTGAAGTAACTGTGTGGCTACAGCTTACGCCCCCTGGGTGAGTACA 4080
Db 4350 GAGACTGGGGGGCGAGACTGTGTGTGTGTGCTGACACTTACCTTCCGGGCTCCGTACCT 4409
Oy 4081 ACCCCCCCAACCAATAGAGAGGTGGCGCTTGGCAGAGGGGAGATCCCTCTTAT 4140
Db 4410 GTGTCCATCTTAACATCGAGAGGTGTGTCTGTCCACACCGAGAGATCCCTTTTAC 4469
Oy 4141 GGGAGGGCGAATTCCTCTGTCTTATCATCAAGGAGAGAAAGATCTTCTTCCATTCA 4200
Db 4470 GCGAAGCTATCCCTCTGAGGTGATCAAGGGGGAGAACATCTCATCTTCTGCCACTCA 4529
Oy 4201 AAGAAAAGTGTGACAGCTTCGGGCGGCCCTTCCGGGTATGTGGCTTGAATCTCAGTGGCA 4260
Db 4530 AAGAAAGTGCACAGCTCCCGGAAAGCTGTGCAATGGGCAATGACCGTGGCC 4589
Oy 4261 TACTACAGGGGTGGACGTCTCCGTAATACCACTCAGGGAGAGTATGGTCTGCGCC 4320
Db 4590 TACTACGGGGTCTTACGCTGTGTGATCTGATCCGACAGCGGCGATGTGTCTGTGTG 4649
Oy 4321 ACCGACGCTCATGACAGGGTATCTGAGGAGCTTGTACTCCGTGATCGACTGCAACGTA 4380
Db 4650 ACCGATGTCTATGATGTGGCTTACCGGCACTTGTGATGTATGATGTGACAGACAG 4709
Oy 4381 GCGGTCACTCAAGTGTGATCTTCAAGTTTGAACCCCACTTCACTCAACACACAGATT 4440
Db 4710 TGTGTACTCAGACATGATTTTCAAGCTTGAACCTTATACATTTAGACACACAG 4769
Oy 4441 GTCCCTCAAGACGCTGTCAAGTACGAGCGCGGGGTGCGACGGGTGAGGGGAGAGCTG 4500
Db 4770 CTCCCGAGATGTCTCTCCAGACTCAACCGCGGAGAGATGTGCGAGGGAGAGCA 4829
Oy 4501 GGCATTTATAGTATGTTTCACTGTGAGCGAGCTCAGAAATGTTTGAACAGTGTAGTG 4560
Db 4830 GGCATCTATAGATTTGTGGACCGGGGAGCGCCCTCCGGGATGTTGATCTGTCTGTC 4889
Oy 4561 CTCTGTGATGTCTACAGCGAGGGCGCATGTGTATGAGTCAACATCTGAGACACACC 4620
Db 4890 CTCTGTGAGTGTATACCGGGCTGTGTGTGTATGAACTCAACGCCCGGAGACTTAA 4949
Oy 4621 GTCAAGGCTCAGGGCGTATTTTCAACAGCGCGGTTTGCCTGTGTGCAAGCACTTTGAG 4680
Db 4950 GTTAGGCTACAGCGTATGAAACACCGGGGCTTCCGTGTGCAAGACATCTTGA 5009
Oy 4681 TTTTGGAGGCAAGTTTTCACCGGCTTACACACATGATGACCACTTCTTCCAAACA 4740
Db 5010 TTTTGGAGGGCGCTTTTACGGGCTCACTCAATATGATGCCACTTTTATCCAGACA 5069
Oy 4741 AAGCAATCGGGGAAAATTTGCAATCTTAAACGCTTACAGGCTTACAGTGTGCGTTAG 4800
Db 5070 AAGCAGATGGGGAGAACTTCTTACCTGTGACCTTACAGGCCACGTTGTGCGTTAG 5129
Oy 4801 GCGAAGCGCCCGCCCGCTGCTGGGAGTCAATGTGAAGTGTGACTGCACTCAAGGCC 4860
Db 5130 GTCTAAGCCCTCTCCCATGTGTGGACAGATGTGAAGTGTGATTCGACTTAAACCC 5189
Oy 4861 ACACTGTGGGCCCAACACTCTCTGTATCCGCTTGGGCTTGTATCAACAGAGTCAAC 4920
Db 5190 ACCCTCATAGGGCAACACCCCTGTATACAGCTTGGGGGTGTTCAGATGAAAGTCAAC 5249
Oy 4921 CTCAACATCTCCGTGACGAATATCATGCGCACTGTGATGACGCACTTGTAGGTCAAG 4980
Db 5250 CTGACGACCACTATCAACAAATCATATGACATGATGTGCGCCACTGTGAGGTCTGTC 5309
Oy 4981 ACCAGCAATGGGTCTTGTGACGGGGAGTCTTGGCGGCGGTGCGGCTTATGCGGCTG 5040
Db 5310 ACAGACACTGGGTGTCTGTGGCGGCTCTGGCTGTCTGTGCGCGATGTGCTGTCA 5369

5041 ACCGGGTGCTTTCATCATCGGCGCTTGACATTTAACGAGGAGCCGTCGTCGCCCG 5100
5370 ACAGGCTGCTGTGTATGAGGAGATGCTCTTGTCCGGGAAGCCGGCAATTATATCT 5429
5101 GACAAAGAGTCTCTATAGAGCTTTTATGATGAGAGAAATGTGCTTAGGCGCGCT 5160
5430 GACAGGAGAGTTCTTACACAGAGTTGATGATGAGAAAGTGTCTCTGACACTTAACG 5489
5161 CTCATTGAAGAGGGGACGAGATAGCCGAGATGCTGAAGTCCAAAGCTTAATTTG 5220
5490 TACATGAGACAGAGGATGATGCTGCTGAGCAGTTCAAGCAGAAAGGCTCTGCTCTG 5549
5221 GACGAGCTTCCAAACAGCTCAAGACATCAACCCACTGTGACGCTTAATAGCCCAAG 5280
5550 CACACCGCTCCGCGCATGAGAGTTATACCCCTGTCTCTCAAGACCAATGGCAGAAA 5609
5281 GTAGAACAAATTCGGGCAAAACATGTGGAATTCATTAGCGGATCAATACCTGCA 5340
5610 CTGAGAGTCTTTGGGGGAAGCACAATGTGGAATTTTCACTAGTGGAAATCAATCTTGGCG 5669
5341 GGAATATCAACACTGCGGAGAAACCTGCAATGCTTCCATGATGCGCTTCACTGCCG 5400
5670 GGCCTGTCAACGCTGCTGTAAACCCGCAATGCTTCAATGATGCTTTTAAAGCTGCC 5729
5401 CTCACAGTCCGCTGTCAACAAACACATATCTTCTCAACATTTTGGGGGCTGGCTA 5460
5730 GTACACAGCCCACTTAACACTGAGCAAAACCTCTCTTCAACATTTGGGGGCTGGCTG 5789
5461 GCATCCCAAAATTCACACACCGCGGGGCACTGAGCTTCTGTGTCAATGAGCTTATGAGGA 5520
5790 GCTGCCAGTCCGCGCGCGCGCGCTGATGCTGCTTGTGGGTGCTGCTGATGCTGCG 5849
5521 GCTCCGTGAGGACATATAGCTTATGATGATGATGATGATGATGATGATGATGATGAT 5580
5850 GCGGCATCGGACAGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 5909
5581 GCGGGGATTTCCGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5640
5910 GCGGGGCTGCGGAGCTCTTGTGACATTAAGATCAATGAGCGGTGAGTCCCTCCACG 5969
5641 GAGGATGCTGCTCAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5700
5970 GAGGACCTGCTCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6029
5701 ATCTGCGGCGCAATCTGCGCGCAACAGCTGAGACCGGGGAGAGCGCGCTGCTCAATGATG 5760
6030 GTCTGCGGAGCAATCTGCGCGCAACAGCTGAGACCGGGGAGAGCGCGCTGCTCAATGATG 6089
5761 AATAGACTCATTTGCTTGTGCTTCAAGAGAAATCACTGCGCGCGCGCGCTGCTCAATGATG 5820
6090 AACCGGCTAATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6149
5821 GAGTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5880
6150 GAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6209
5881 CTCAAGAACTCAACAACTGATTAATGAGGACTGCGCGCAATCCCATGCGCGCTGCTG 5940
6210 CTGAGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6269
5941 CTCGCGATGCTGAGGAGCTGAGGATGAGCACTCAATCAAGCTTAAATTTGCTGAC 6000
6270 CTAAAGGACATCTGAGGAGCTGAGGATGAGGAGCTGAGGAGCTTAAAGCTTGAATG 6329
6001 TCCAAATTAATTCCAAAGATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6060
6330 GCGAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6389
6061 GCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6120
6390 GGGGTCTGCGGAGAGAGCGGCAATTAATGACATCTGCTGCGCACTGTGAGGCTGATGATCT 6449

6121 GGCATGTCGCTTGGGCTCATAGAAATCAAGGAGCTTAAGACTGATGAATATCTG 6180
6450 GGAATGTCAAAAACGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 6509
6181 CAGGAGACTTCTCTATCAATGTTTACAGGAGGAGCAATGCTGCGGAAACCGCGCA 6240
6510 AGTGGAGCTTCCCAATTAACGCTTACACAGGAGCCCTGTACTCCCTTCTGCGCGC 6569
6241 AACTTAAAGTCCCATCTGAGAGGCTGCGGCTCAGAGTACCGCGAGAGTACGACGAC 6300
6570 AACTTAAAGTCCCATCTGAGAGGCTGCGGCTCAGAGTACCGCGAGTACGAGATGAGG 6629
6301 GAGTCAATACCACTAATACAGAGCTCAACCATGATTAATTTGAAAGTCCCTGCACTA 6360
6630 GAGGACTTCCCACTGATGAGGATGATGATGATGATGATGATGATGATGATGATGATG 6689
6361 CCTCTCCGAGTCTTCTTCTGAGGAGAGTCAAGTCAATGATGATGATGATGATGATG 6420
6690 CCATGCGCGAATTTTACAGAAATGACGAGGAGTCCCTCAACAGTGTGCGCCCT 6749
6421 CCGAAGCCGTTTTCCGAGTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6480
6750 TGCAGCCCTTGTGCGGAGAGGATCAATTAAGATGAGTCAACAGTACCGGCTG 6809
6481 GGGTCCAGGCTTCTTGCACCTGAACCCGACACAGACGATATTGATGCTCATACA 6540
6810 GAGTCCCAATTAATCTGAGAGCCGAGACGAGTACGCTGTTGATGCTCATGCTCAT 6869
6541 GATCATCTTCAATCAACGAGGAGATGCTGAGCGCGGCTTTAGCGCGGGGTCAACCCCA 6600
6870 GATCCCTCCCAATTAACAGAGAGCGCGCGGAGAAAGGTTGCGAGAGGATCAACCCCT 6929
6601 TCCGAGGCAAGCTCTTCCGAGAGGAGTATGAGGACATATGCTGCGAGCACTGAGC 6660
6930 TCTATGCGAGCTCTGCTGAGGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6989
6661 ACCACGCGCAAGCTATGATGAGTCAATGATGATGATGATGATGATGATGATGATGATG 6708
6990 GCCAACCATATCTCCCTGAGCGCGGAGCTCATAGAGCTTACCTCTGTGAGGAGAG 7049
6709 ATGGGGGCGAGTGAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 6768
7050 ATGGGCGGCAATCAACAGAGGTTGATGATGATGATGATGATGATGATGATGATGATG 7109
6769 GACCCATGCTGCAAGAAAGAGGAGCTTTGAGCTTCAATCAATCAAGATGATGATG 6828
7110 GATCCCTGCTGCAAGAGGAGTGAAGGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTG 7169
6829 CCAAGAGAGGTTCCCAACGCTTACCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6888
7170 AAGTCTGAGATTTCCGCGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7229
6889 CTGTGGAATCTGAGAAAGGAGGATTAACAAACCGGCACTGTTGCGGCTGCTGCTGCTG 6948
7230 CTAGTAGAGAGCTGAGAAAGCTGATCAAGAACACTGTGCTGCTGCTGCTGCTGCTG 7289
6949 CTTCTCTTGAAGAAACCCGAGGCTTCCGCAAGAGGAGGAGGAGGAGGAGGAGGAGG 7008
7290 CCACTTCCAGAGTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7349
7009 GAGGATCTCAATGAGATGCTTCAACAGCTGAGCTTAAATGATGCTTGTGCGAGCCGCC 7068
7350 GAATCAACCTTATCTACTGCTTGGCGAGCTTGCACCAAAAGTTTGGCAGCTCTCA 7409
7069 CCAAGCGGCAATTCAGGCTTTTCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7128
7410 ACTTCCGCGATTAACGAGGAGCAATTAACAGCAATCTCTGAGGCGCGC----- 7457
7129 CTTCTGATGAGTGGCGCTTGTGAGAGACAGGTTTCAATCTTCAATGAGGCGCGCTGAG 7188
7458 CTTTCTGAGCTGCGCGCGAGCTGAGGTTGATGATGATGATGATGATGATGATGATGATG 7517
7189 GGGAGCTTGAAGATCAGACTGAGGCTGAGCAGGATGAGGCCCAACCCCGCGAG 7248


```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/02225
FILING DATE: 19910329
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CIOTTI, THOMAS E.
REGISTRATION NUMBER: 21,013
REFERENCE/DOCKET NUMBER: 2300-0101.44
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 327-7250
TELEFAX: (415) 327-2951
TELEX: 706141
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 9401 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
PCT-US91-02225-9

Query Match      45.3%; Score 4118; DB 5; Length 9401;
Best Local Similarity 66.2%; Pred. No. 0;
Matches 6029; Conservative 20; Mismatches 2971; Indels 90; Gaps 5;

QY 1 ATGAGCAAAATCTTAAACCTCAAGAAAGAAAGAAACCAACCAACCGTCCGCCA 60
DB 342 ATAGACAGAAATCTTAACCTCAAAARAAACAAAGTAACACCAACGTCGCCACAG 401
QY 61 GACGTAAAGTTTCGGGCGGCGGCAAGATGTTGGCGAGATATCTTGTTCGCCGAG 120
DB 402 GACGTAAAGTTTCGGGCGGCGGCAAGATGTTGGCGAGATATCTTGTTCGCCGAG 461
QY 121 GGGCCAGGTTGGGTGTGGCGCGCAAGAAAGTAAGAGGTTTCCAGCGGTGCGAG 180
DB 462 GGGCCAGGTTGGGTGTGGCGCGCAAGAAAGTAAGAGGTTTCCAGCGGTGCGAG 521
QY 181 AGGCGCAGGCGCAATCCCTAAGATCGGCGCTCCACTGGGAAATCTGGGAAACAGCA 240
DB 522 AGAGCTAGGCTTAATCCCAAGGCTCTGCGGCCGAGGCGAGACCTGGGCTCAGCGCG 581
QY 241 TACCCCTGCGCCCTATATCGGAAATGAGGAGCTGCGCTGGCGAGAGTGGCTCTG 300
DB 582 TACCTTGGCCCTCTATAGGCAATGAGGCTGGGCGGTGGGCGGATGGCTCTGCTCC 641
QY 301 CGAGTTCCGCTCTCTTGGGCGCCCAATGATCCCGGCAATAGGTGCGCAAGTGGT 360
DB 642 CGTGGCTCTGCGCTAGCTGGGCGCCCAAGACCCCGCGGTGCGCAATTTGGGT 701
QY 361 AAGGTATGATATCCCTTAAGTGGGCTTGGCGAAGCTTATGGGATCATCCCTGTG 420
DB 702 AAGGTATGATATCCCTTAAGTGGGCTTGGCGAAGCTTATGGGATCATCCCTGTG 761
QY 421 GGGCGCCGCTCGCGCGGCTGCGCAGAGCTCTGCGCATGAGCTGAGAGTCTGAGAG 480
DB 762 GGGCGCCGCTCTTGAAGGCGCTGCGCAGAGGCTTGGCGCATGAGCTGAGAG 821
QY 481 GGGGTAATTTTGAAGAGGAACTTACCGGTTGCTCTTTTCTATCTTTGCTGGC 540
DB 822 GGGGTAATTTTGAAGAGGAACTTACCGGTTGCTCTTTTCTATCTTTGCTGGC 881
QY 541 CTGCTGCTGATCAACACCGGCTCTGCGCTGCGAGTGAAGAAATCAAGTACGCG 600
DB 882 CTGCTGCTGATCAACACCGGCTCTGCGCTGCGAGTGAAGAAATCAAGTACGCG 941
QY 601 TACATGATGATCAAGCTGCAACCAATGACAGATTAAGTGAAGTCTGAGCTGTC 660
DB 942 TACATGATGATCAAGCTGCAACCAATGACAGATTAAGTGAAGTCTGAGCTGTC 1001
```

```
QY 661 CTCACGTCCCGGGGCGGTCGCCGCGGAGAAAGTGGGGAATGATCTGAGTGGATA 720
DB 1002 CTCACGTCCCGGGGCGGTCGCCGCGGAGAAAGTGGGGAATGATCTGAGTGGATA 1061
QY 721 CCGGCTCAACCGAATGAGCGGCGGCGGCGGCGGCTCAAGAGAGGCTTGGCGAG 780
DB 1062 CCGGCTCAACCGAATGAGCGGCGGCGGCGGCGGCGGCTCAAGAGAGGCTTGGCGAG 1121
QY 781 CACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 1122 CACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1181
QY 841 TGGGCGGCGGTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 1182 TGGGCGGCGGTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1241
QY 901 GTCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 1242 AGCGAAGTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1301
QY 961 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 1302 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1361
QY 1021 CCGGAGTCAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1362 CCGGAGTCAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1421
QY 1081 TACTTCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
DB 1422 TACTTCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1481
QY 1141 GTGAGCGCGGCAACCAATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1482 GTGAGCGCGGCAACCAATGATGATGATGATGATGATGATGATGATGATGAT 1541
QY 1201 AGCTTATTTGACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
DB 1542 AGCTTATTTGACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1601
QY 1261 CACATCAACGCGACCGGCTGAACTGAACTGAACTGAACTGAACTGAACTGAACT 1320
DB 1602 CACATCAACGCGACCGGCTGAACTGAACTGAACTGAACTGAACTGAACTGAACT 1661
QY 1321 CTGTTCTACACCAAGCTTCAATCTGTAAGATGATGATGATGATGATGATGAT 1380
DB 1662 CTGTTCTACACCAAGCTTCAATCTGTAAGATGATGATGATGATGATGATGAT 1721
QY 1381 AGTATGAGGCGCTTCCGGGTGGGATGGGCGGCTTGCATATGAGGATATGAT 1440
DB 1722 CCGCTTAACGATTTTGAACAGGAGGCTGAGGCTTATGATGATGATGATGAT 1779
QY 1441 CCAGAGATATGAGACCTTATGCTGAGCACTACCAAGAGGAGTGGCTGTCTCC 1500
DB 1780 ----CCGACCAAGGCGGCTTATGCTGAGCACTACCAAGAGGAGTGGCTGTCTCC 1835
QY 1501 GCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
DB 1836 GCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1895
QY 1561 ACCGAGAGGCTTGGAGCGGCACTTACAGTGGGAGGAGATGAGCAGATGCTTCTA 1620
DB 1896 ACCGAGAGGCTTGGAGCGGCACTTACAGTGGGAGGAGATGAGCAGATGCTTCTA 1955
QY 1621 TTGAACAGCACTGACACCGCTGAGGATGATGATGATGATGATGATGATGATGAT 1680
DB 1956 CTTAACAATATCAAGGCAACCGCTGAGGATGATGATGATGATGATGATGATGATGAT 2015
QY 1681 GGTACACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 2016 GGTACACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2073
```

QY	1741	A	CGAC	CTGTTG	TG	CCCA	CGGAC	CTGTTT	AG	GA	AGCA	TCTCG	TAA	CTACT	ACT	CAAA	1800						
Db	2074	---	CC	TGCA	CTG	CCCA	CTG	ATTG	CTT	CCG	AA	GA	TCC	GA	CGCA	CTT	CCG	2129					
QY	1801	T	CGG	CTC	TG	GG	CC	CTG	CTCA	CG	CA	AG	TG	CG	AT	CA	CGG	CTC	TG	1860			
Db	2130	T	CGG	CTC	CG	GG	CTC	TG	MT	CA	CA	CG	AG	TG	CG	AT	CA	CGG	CTT	TG	2189		
QY	1861	C	ATT	A	CC	CT	GC	A	CGT	T	AA	CT	TA	TA	CC	TA	TT	AG	AG	GG	TT	1920	
Db	2190	C	ATT	A	CC	CT	TG	A	CCAT	CA	CT	AA	CA	TA	TT	AA	TA	CA	GA	TG	TG	2249	
QY	1921	G	AG	CA	CA	GG	CT	CA	CGG	CT	GA	T	GA	TT	CA	CT	CG	TG	GG	AG	TC	1980	
Db	2250	G	A	CA	CA	GG	CT	GA	AG	CT	CG	CA	CT	GA	AG	CG	GG	CA	AG	CT	TG	2309	
QY	1981	A	G	A	G	A	G	A	G	T	C	A	G	T	C	T	T	T	G	T	G	A	2040
Db	2310	A	G	G	A	C	A	G	T	C	C	A	G	T	C	T	T	T	G	T	G	A	2369
QY	2041	T	G	C	T	T	A	C	T	G	G	A	C	T	T	G	A	C	T	G	A	2100	
Db	2370	T	G	T	T	C	T	T	C	A	C	A	C	C	T	T	A	C	T	G	A	2429	
QY	2101	G	T	G	A	C	G	T	A	C	T	T	A	T	A	T	A	T	A	T	A	2160	
Db	2430	G	T	G	A	C	G	T	A	C	T	T	A	T	A	T	A	T	A	T	A	2489	
QY	2161	G	A	G	T	G	G	A	T	A	C	T	T	A	T	A	T	A	T	A	T	2220	
Db	2480	G	A	G	T	A	C	G	T	T	C	T	T	A	T	A	T	A	T	A	T	2549	
QY	2221	A	T	G	C	T	A	T	T	G	G	A	C	G	A	C	G	A	C	T	G	2280	
Db	2550	A	T	G	A	T	G	C	T	A	T	T	G	G	A	C	G	A	C	T	G	2609	
QY	2281	G	G	A	G	C	G	A	C	T	A	T	T	G	T	G	A	C	T	T	2340		
Db	2610	G	A	T	G	C	T	G	A	C	T	T	A	T	T	G	T	G	A	C	2669		
QY	2341	T	A	C	A	T	A	C	A	G	G	T	A	G	T	A	C	T	A	T	2400		
Db	2670	T	A	T	T	G	A	A	G	G	T	A	G	T	A	C	T	A	T	A	2729		
QY	2401	A	G	C	T	A	T	G	C	T	T	A	G	A	C	T	T	A	G	A	2460		
Db	2730	C	T	C	T	G	C	T	T	G	C	G	T	G	C	C	A	G	T	G	2789		
QY	2461	C	A	G	A	T	A	G	A	C	G	C	T	G	T	G	T	A	T	A	2520		
Db	2790	T	C	G	T	G	T	G	G	T	G	T	T	C	T	C	G	T	G	T	2849		
QY	2521	A	C	C	T	T	C	A	G	C	G	G	T	T	T	T	T	T	T	T	2580		
Db	2850	C	G	C	T	A	T	A	C	A	G	T	G	G	C	T	T	T	T	T	2909		
QY	2581	G	T	C	A	G	A	G	T	G	G	A	C	T	T	A	G	A	G	T	2640		
Db	2910	C	T	G	A	C	G	T	G	A	T	T	C	C	C	C	C	T	A	C	2969		
QY	2641	G	T	C	A	T	A	T	T	C	A	C	C	A	G	T	G	T	T	A	2700		
Db	2970	A																					

Db	3150	GTATCATTTAAGTTAGGGGGGCTTACTGGACCTATGTTTATAACCATCTCACCTCCTT	3209
Qy	2881	TCGGATTGGGCTGCTAGTGGCCCTGGCGGACCTGGCCGGTGCCTGTTAGCCTTATCATCTTC	2940
Db	3210	CGGAGCTGGGGCGCACACGGCTTGGCAGATCTGGCCGTGGCTGTAGAGCCAGTCGTCTTC	3269
Qy	2941	AGTCCCATGGAGAAAGAAAGTCATTGTTCTGGGGAGCCGAGACAGCTGTTGTGGGACATT	3000
Db	3270	TCCCAATATGAGAACCAAGCTCATACGTGGGGGGGCAATACCGCCGGGTGGGTGATCATC	3329
Qy	3001	TTACACGAGACTCCCGTGTCCGCCGCCGACTTGTGTGGAGAGTCTCTTGGCCCACTGAT	3060
Db	3330	ATCAACGGCTTGCCTGTTTCCGCCCGAGGGGCCGGAGATACCTGCTCGGGCCAGCCGAT	3389
Qy	3061	GGCTATACCTTCCAAAGGGGTGAGTCTTCTGCCCCCATCACTGCTTACGCCCGACAGACA	3120
Db	3390	GGATATGCTTCCAAAGGGGTGAGAGTTCTGGCCGCCCATCACCGGGTACGCCGACAGACA	3449
Qy	3121	CGTGCCCTTTTGGGACCATATAGTGTAGCATGACGGGGGCCGCAAGAACAAAGAGCT	3180
Db	3450	AGGGGCTCTCCATGGGTGCATATACCAAGCTTAACTGGCCGGGACAAAAAACAAAGTGAG	3509
Qy	3181	GGGGAATTTGAGCTCGTCCACAGTCACATCAGTCCCTTCCTCGGAAATCATCATCTCGGGG	3240
Db	3510	GGTAGGTCAGATTGTGTACATCTGCTGCCCAACCTTCTTGGCAACGTGATCATATGGG	3569
Qy	3241	GTTTTGTGACTGTCTACATGAGACTGGGCAACAAAGACTCTGGCCGGCTCACGGGGTCCG	3300
Db	3570	GTTGTGCTGAGCTGTACACACGGGGCCGGAACAGACCATCGGTCAACCAAGGGCTCT	3629
Qy	3301	GTCACGAGATGTATCTCAAGTCTGAGAGGGGACTTAATAGGGTGGCCAGCCCCCTGGG	3360
Db	3630	GTCATCCAGATGTATCAATGTATGACCAAGACTCTTGCGGCTGGCCCCGCTTGCAGAGGT	3689
Qy	3361	ACTAAATCTTTGAGACCGGTGACGCTGTGTGAGCCGGTGCACCTGATACCTGTACACGGGAC	3420
Db	3690	ASCCGCTATTGACACCTGTGACTTGGGCTCTCGGACTTTTACTGTGTACGAGGAC	3749
Qy	3421	GCTGATGTCAATCCCGGCTCGAAGACGCGGGGACAAACGGGGAGCGCTACTCTCCCGAGA	3480
Db	3750	GGCATGTCAATCCCGTGGCGCGGGTGTGATAGACGGGGACAGCTGTGTGCCCCGG	3809
Qy	3481	CTCTTTTCCACTTTGAAGGGGTCTTCAAGAGACCCGGTGTATGCCCCAGGGCCAGCT	3540
Db	3810	CCCATTTCTTACTTGAAGGCTCTCCGGGGGGTCCGCTGTGTGTGCCCCCGGGCAAGCC	3869
Qy	3541	GTCGAGCTTCCCGGACGCTGTGTGCTTCGGGGCGGTGTAAGTCATATGATTTATC	3600
Db	3870	GTGGGCATATTTAAGGCGCGGGTGTGACCCGTGTGATGTGCTAAGCGGTGAGATTATC	3929
Qy	3601	CCCGTTAGACATCTGCATCTGTACCCGGTCCCCCACTTTAGTGAACAAGCACACA	3660
Db	3930	CTGTGTGAGAACCTTAGAGACACCATATAGGTCCCCGGGTGTTCACGGATATCTCTTCCA	3989
Qy	3661	CCTGCTGTGCCCGACCTATCAGTGGGGTACTTTCATGCCCCGACCTGCAGGTGGAAG	3720
Db	3990	CCAGATGAGCCCCAGAGGTTCCAGATGGCTTCACTTCATCTCCACAGGCAGCGGCAAA	4049
Qy	3721	AGCACCAAAAGTTCCTGTGCGATTATGTCTGTCTCAGGGGTATTAAGTGTAGTCTTAATCCC	3780
Db	4050	AGCACCAAGGTCGGGCTGCAATAGCAGCTCAAGGGCTATTAAGGTGTAGTACTCAACCCC	4109
Qy	3781	TCAGTGGCTGCACCTCGGGGTTTGGGGCGTATTTGTCTAAGGACATGGCATCAATCCC	3840
Db	4110	TCGTGTGTGCAACACTGGGCTTTGGTGTCTTACATGTCTCAAGGCTCATGGATGATCTYT	4169
Qy	3841	AAACATTAGACTGTGAGTCAAGACTGTGACGACCGGGGCGCCATCACGTACTCCACATAT	3900
Db	4170	AAACATTAGGACCGGGGTGTAGAAACAATTACACTGGACGCCCATCACTACTCCACTAC	4229
Qy	3901	GGCAAAATTCCTGCGCGATGGGGGCTGTGTGGGGCGGCGCTTACGACATCATCATATGTGAT	3960

D	4230	GGCAAGTTCTTGGCCGACGGGGGTGCTGGGGGGCGCTTATGACATTAATAATTGTGAC	4289
Q	3961	GAATGCCATGCGCGAGACTCTTACCAACCATCTTGAGCATCGGAACAGTCTTGATCAAGA	4020
D	4290	GAGTGGCACTTCAAGATGGCCATCATCTTTGGKCATCGGCACTGTCTTGACCAAGCA	4349
Q	4021	GAGACAGCTGGGTCACACTTACTGTGTGGCTACAGTACGCCCCCTGGGTCAGTGACA	4080
D	4350	GAGACTCGGGGGGAGACTGGTGTGTGTCCGACACCGCACCCCTCGGGCTTCGTCACT	4409
Q	4081	ACCCCCACCCCAATPAGAGAGGTGGCCCTTGGGACGAGGGCGAGATCCCTTCAT	4140
D	4410	GTGCCCCATCCCAACATCGAGAGGTGTCTGTTCACACACCGAGAAATCCCTTTTAC	4469
Q	4141	GGAGGGCGATTCCTCCGTCTTATCATCAAGGAGAGAACATCTGATCTTCTGCAATTCA	4200
D	4470	GGCAAGGCTATCCCTCTGAAAGTATCAAGGGGGGAGACATCTATCTTCTGTCAATTCA	4529
Q	4201	AAGAAAAAGTGTACGAGCTCGCGCGGCCCTTGCGGGTATGGGCTTGAATCAAGTGCA	4260
D	4530	AAGAAAGAGTGTGACGAACATCGCCGCAAGCTGTGTGCATTTGGGCATCAATGCGTGCC	4589
Q	4261	TACTACAGAGGTGTGAGCTCTCCGTATATCAACTCAGGGAGACGTATGCTGTGCC	4320
D	4590	TACTACCGCGGTCTTGAAGTGTCCGTATCCGACACAGCCGAGATGTGTGTGTGTGCA	4649
Q	4321	ACCGACGCCCTCATGACAGGGTATATCTGGGAGACTTGTGACTCGATGTGACTGCAACGTA	4380
D	4550	ACCGATGCCCTCATGACCGGCTATACCGGCGACTTGCATCGGTGATAGACTTTCATACG	4709
Q	4381	GCGGTCACTCAAGTTGTAGACTTCAAGTTAGACCCCAATTCACCATTAACCAACAAGATT	4440
D	4710	TGTGTACACCGACAGTGTGATTTTACGCTTGACCTTACCTTACATGAGACATCAACG	4769
Q	4441	GTCCCTTAAGAAGCTGTGTCTACAGTACGAGCGCCGGGGTCCGACGGGTATGGGGAAAGCTG	4500
D	4770	CTCCCCAGAGATGTGTCTCCCGCACTCAAGTGTGGGCGAGACTGTGCGGGGAGAGCA	4829
Q	4501	GGCATTTATAGTATGTTTCCACTGTGGAGAGGAGCTCAGGAATGTTTGAAGTGTAGTG	4560
D	4830	GGCATTTACAGATTTTGTGTGACCGGGGAGGCGCCCTTCGGCATGTTCGATCGTCCGT	4889
Q	4561	CTCTGTGAGTGTCTACGAGCAGGGGCGCGCATGTATGAGCTCACACCATCGGAGACACC	4620
D	4890	CTCTGTGAGTGTATGACGAGCGGTGTCTTGTATGAGCTCACGCGCGGAGACTACA	4949
Q	4621	GTCAAGGCTCAGAGCGGTATTTCAACAACGCGCGGTTTGGCTGTGTGCGCAAGACATTTGAG	4680
D	4950	GTTTAGGCTACAGAGGTACATGAMACACCCGGGGCTTCCCTGTGTCCAGGACATCTTGAA	5009
Q	4681	TTTTGGGAGAGAGTTTTCACCGGCGCTCAACAATAGATGGCCACTTCTTCCCAACA	4740
D	5010	TTTTGGGAGGGCGTCTTTACAGGCGCTCACTATATAGATGCCACTTCTATCCAGACA	5069
Q	4741	AAGCAATCGGGGAAAAATTTGCACTACTTAAACAGCTTACAGGCTTACAGTGTGCGTAGG	4800
D	5070	AAGAGAGTGGGGAGAACCTTCTTAACTGTGATAGCATCAAGACCAACGTGTGGCTAGG	5129
Q	4801	GCCAAAGCCCCCCCCGCTCTGTGAGCGTACATGTGAAAGTGTGTGACTGCACTCAAGCC	4860
D	5130	GCTCAAGCCCCCTCCCATCGTGGGAGACAGATGTGAAAGTGTGTGAATTCGCTCAAGCCC	5189
Q	4861	ACACTGTGGGGCCCCACAACCTCTCTGTACCGCTTGGGCTGTGTTACCAAGAGGTACC	4920
D	5190	ACCTTCATGGGCGCAACACCCCTGTATACAGACTGGCGCTGTTCAGAAATGAATCAC	5249
Q	4921	CTCACACATCCCGTGACGAAATATACCTGCGCACTTGACGAAAGCGCACTTGAAGTCAATG	4980
D	5250	CTGACGACCAAGTCAACCAATATATATATGACATGCAATGGGCGGACCTTGAAGTGTGCT	5309
Q	4981	ACCAACACATGGGTCTTGTGGCAGGGGAGTCTTGGCGGCGTTCGCGCGGTATTTGCTGGCG	5040
D	5310	ACGAGACACTGGGTCTCTGTGTGGCGGCTCTCCGTGCTCTTGTGGCGCGCTATATGCTGTGCA	5369

OY	5041	ACCGGGTGTGTTTGATCATTCGCGCGCTTGACATTAACGAGGCGCGTCTGGCGCG	5100
Db	5370	ACAGGCTCCGGTCTATAGTGGGAGGGTGTCTTTGTCGGGAAGCCGGCATATATACCT	5429
OY	5101	GACAAAGAGGTCTCTTATGAGGCTTTTGTATGAGATGAGAAATGTGCTCTAGGCGGCT	5160
Db	5430	GACAGGGAAGTCCCTACCGAGAGTTGATGAGATGGAAGAGTGCTCTCAGCACTTACCG	5489
OY	5161	CTCATTTAAGAGGGGGCAGCGGATATGCGGAATGTCTGAATCTCAAGATTCGAAGCTTATTG	5220
Db	5490	TACATTCAGCAAGGAGATGATGTCTCGCCGACAGTTCAAGAGAAAGGCGCTCGGCTCTG	5549
OY	5221	CAGAGAGCTTCCAAACAAGCTCAAGACATACACCCACTGCGAGGCTTCAATGSCCAAG	5280
Db	5550	CAGACCGCGTCCCTCAGCGAGAGTTATGCGCCCTGCTGTCCAGACCACTGGCAAAAA	5609
OY	5281	GTAGAACAATTTCTGGGCCAAACACATGTGGAATTCTTAAGCGGCATCAATACCTCGCA	5340
Db	5610	CTCAGAGACCTTCTGGGCGAAGACATATGTGGAATTCATCAGTGGGATACATATCTTGGCG	5669
OY	5341	GGACTATCAACACTGCGCAGAGGGAACCTGCGAGTACCTCAATGAAGGGGTTCAAGTCCGCC	5400
Db	5670	GGCTTGTCMAAGCTGCGCTGTGTAACCCCGCCATGTCTTCAATGATGGCTTTTACAGTGTCT	5729
OY	5401	CTCACAGTCCGCTGTCAACAAGACCACTATCTTCTCAACATTTTGGGGGCGTGCCTA	5460
Db	5730	GTCACAGAGCCACTATACCATAGCCAAACCTCTCTTCAACATATTGGGGGGGTGGGTG	5789
OY	5461	GCATCCCAATTTGACACACCCCGGGGGGCCACTGAGCTTCTGTCAAGTGGCTTATGTGGGA	5520
Db	5790	GCTCCCAAGCTGCGCGGCCCGGGGCGGCTATGCTCTTTGTGGGCGGTGGCTTAAGTGGC	5849
OY	5521	GCTCCCGTAGGCACTATATAGCTTATGATGAAGTGTCTTATGACACTCTGGCAGGATATAGT	5580
Db	5850	GCCGCCATCGGCACTGTGTGAATGGGGAAGTCTCTATAGACATCTTTCAGAGGTTATGGC	5909
OY	5581	GCGGGCAATTTGCGGGGCTCTGTGCAATCAAGATCTGTGCGAGAAAGCCTTCATG	5640
Db	5910	GCGGGCGTGGCGGAGCTTTGTGGCAATTCAGATCATGAGCGGTGAGTGTCCCTCCACG	5969
OY	5641	GAGATCTCCGTCACCTTGCTGCGCCGGAATTCGTCTCCGGGTGCTTGATAGTGGAGTCT	5700
Db	5970	GRGAACCTGTGCATCTACTGTCCCGCCATCTCTGCGCCGAGGCTCTGTATGTGGCGTGT	6029
OY	5701	ATTCGCGCGGCATTTCTGCGCCGACAGCTGGGAGACGGGGGGAAGGCGCGTTCATGTAGT	5760
Db	6030	GTCGTGCAAGCAATCTGCGCGGCAAGTTGGCCCGGGCGAGGGGCGCATGTGATG	6089
OY	5761	AATAGACTCATTTGCTTTGCTTCCAGAGAAATTAAGTCGCGCCACCCACTTACTGTAGC	5820
Db	6090	AACGGGCTGATAGGCTTTCGCTCCCGGGGGAACATTTTCCCAAGCACTTACGTGCGG	6149
OY	5821	GAGTCGAGTGTGCGAGCGGTGACCACTACTTGGTCTCCCTTACATTAACAGCTGT	5880
Db	6150	GAGAGCGATGTAGCTGCCCGCGGTACTGTCCATCMGCMGCAAGCTCATGTATACCACTCTC	6209
OY	5881	CTCAGAAAGACTCAACACTGATTTACTGAGAGCTGCCCATCCCATGCGGCGCTCGTG	5940
Db	6210	CTGAGGCGACTGACCAAGTGGATTAAGCTCGGAGGTATCACTCATCTCCGATTTCTGTG	6269
OY	5941	CTCGGCATGTGTGGGACTGGGTTTGGACCACTCTTAACAAGCTTAAATAATTTGCTGAC	6000
Db	6270	CTAAGGACATCTTGGAGTGGATATGCGAGGTGTGTGACGACHTTAAAGACTTGGCTAATA	6329
OY	6001	TCCAAATTAATCCCAAGATGCGCGGCTCCCTTTTGTCTCTGTCTCAAAAGGGGTACAG	6060
Db	6330	GCTAAGCTCATGCGACAGCTGCGCTGGGATCCCTTTGTGTCCTCCAGCGCGGGTATTAAG	6389
OY	6061	GGCGTGTGGCGCGGCACTGGATATATACACACAGGTGTCTTGGCGCGCAATATCTCT	6120
Db	6390	GGGGTCTGGCGAGGACGCGCATATACACTGTGCTGCAACTGTGGAAGCTGMAATACCT	6449

QY 6121 GGCAGATGTCGGCTGGCTCAGATGAATACAGGGGCTTAAGACCTGATGATATCTGG 6180
Db 6450 GGACATGTCTCAAAAAAGGGACGATGAGATCGTCGGTCTTAGAGACTGGAGAAATCATGTGG 6509
QY 6181 CAGGGGACCTTTCTTATGATTTGTTTACACGAGGGCCAGTGCCTGGCAAAACCCGGCCA 6240
Db 6510 AGTGGAGCTTCCCTATTAATGCTACACAGGGCCCTGTATCCCTTCTGCGCG 6569
QY 6241 AACTTAAAGTCGCCATCTGAGAGGTGGCGCTCAGAGTACGGGAGAGTACCGACAGAC 6300
Db 6570 AACTACAGTTCGGCTATGAGAGGTGTCTGACAGAAATATGTGAGATTAAGGCAAGTGTG 6629
QY 6301 GGGTATACCACTATATACAGAGACTACCACTGATTAACCTTAAAGTCCCTGCCAACTA 6360
Db 6630 GGGGACTTCCACTAGTACGGGTATGACTACTGCAATCTCAATATGCTCCGTGACAGTTC 6689
QY 6361 CCTCTCCGAGTCTTTTCTCGGTGAGACGGAGTGCAGATCCATAGTTGGCCCCACA 6420
Db 6690 CCATGCCCCGAAATTTTTCACAGAAATTGAGACGGGGTGCCTTACATAGTTGGCCCCCC 6749
QY 6421 CCGAAGCGTTTTTCGGGATGAGGTCTGTTCTGCTTGGGCTTAAATTCATTTGTGCTC 6480
Db 6750 TGCAAGCCCTTGCTGGGGAGGAGGTATCATTAAGATAGGACTCCAGAAATACCCGGTA 6809
QY 6481 GGGTCCAGCTTCTTGGGACCTTGAACCCGACACAGAGTATGATGCTCATGCTAAC 6540
Db 6810 GGGTGCATTAATCTTGGAGCCGACCGACGTGGCCGTGTGACGTCCATCTCCTCACT 6869
QY 6541 GATCATCTCATATACGGCGGAGACTGACAGCGGCGGTTTAAAGCGGGGGTTCACCCCA 6600
Db 6870 GATCCCTCCCATATACAGACAGAGGGGCGCGGCGAAGGTTTGGCGAGGAGATCACCCCC 6929
QY 6601 TCCGAGGAGAGTCTCTCGGCGAGGAGGACTATCGGACCATGCTGCGAGCCACTGACAC 6660
Db 6930 TCTGTGGCAGGTCTCTCGGCTTAGCCAGCTATCCGTTCATCTCAAGGCAATTTGACCC 6989
QY 6661 ACCACGGCAAAAGCTATGATGTGACATGTGTGACTTAACTT-----GTTT 6708
Db 6990 GCTAACATGACTCCCTGATGCTGAGCTCATAGAGCCAACTCTTATGAGGACAGAG 7049
QY 6709 ATGGGGGGCGATGTAATCTCGATAGATCTGGGTCAAGATGTCGTTCTGAGACTCTCTC 6768
Db 7050 ATGGGCGGACATACACAGAGGTATGATCAAAAAAAGATGTGATTTCTGACCTCTTC 7109
QY 6769 GACCAATGTCGAGAAAGAGAGGACGACTTGGACCTTGCATACCATAGAAATAGATGTC 6828
Db 7110 GATCGGCTTGTGGGAGAGAGAGGAGGAGATCTCGTACCCGAGAAATCTGTGG 7169
QY 6829 CCGAAGAAAGAGTCCACAGCTTTACCGGCTTGGGCAAGGCTGATTAACCCACCG 6888
Db 7170 AAGTCTCGAGATTTGCCAGGCGCTGCGCTTGTGGGCGGCGCGGAGCTATTAACCCCG 7229
QY 6889 CTGTGGAATGCTGAGAAAGGCGAGATTACCAACCGGCGCACTGTGTGGGGGCTGTGCTC 6948
Db 7230 CTATGTGAGACGTGAGAAAGGCGGAGCTACGAAACCACTGTGTTCATAGGCTGTCCGTT 7289
QY 6949 CCTCTCTGAGAAACCCCGAGGCTTCCCAAGAGAGCGCGGACAGTGGGCTTAAGT 7008
Db 7290 CCACCTCCAAAGTCCCTCTGTGTGCTCGGCTCGGAAAGAGCGGAGCGTGTCTCACT 7349
QY 7009 GAGGACTCCATAGAGATGCTCTTCAACAGCTGCGCATTAAGTCTTTGGCCAGGCCCCC 7068
Db 7350 GAATCAACCTTATCTACTGCTTGGCCGAGCTCGGACAGAAAGTTTGGGAGCTCTCA 7409
QY 7069 CCAAGGGGAGATTCAAGGCTTTTCCAGGGGGGCGGGGCGTGCAGATTCCGGGAGTCAAGG 7128
Db 7410 ACTTCGGGATTAAGGGGCAAAATACGAACAATCTCTGAGGCC----- 7457
QY 7129 CTTCTCTGATGATGGCTTTTGGAGACAGGTTTCATCTTCCATGCCCCCTCTGAG 7188
Db 7458 CTTTGTGGCTGCCCCCGGAGCTCGAGCGTGAAGTCTTWTCTCTCATGCCCCCTCTGAG 7517
QY 7189 GGGAGCTTTGAGATTCAGACTGAGCGCTGAGAGGTAGAGCCCCCAACCCCCCCAG 7248

Db 7518 GGGAGCTTGGGATTCGGA----- 7537
QY 7249 GGGGGGGTGGACAGCTCCGGCTCGAGCTGGGGTCTGGTCTTACTTGTCTCGAGAGAGAC 7308
Db 7538 -----TTTAGCAGCGGTATATGTTCAACGCTACGTAAGTGAAGCCAGCGG 7583
QY 7309 GACTCGGTGTGTCTGCTTCATGTATATCTCTGAGACGGGGCTTAAATCTCTTGT 7368
Db 7584 GAGGATGTGTGTGCTCAATGTCTTACTTGTGACAGGCGCACTGTCACCCGTGC 7643
QY 7369 AGTCCGAAGAGAAAGTATCCGATTAACCTTTAGAGCACTCCCTGTGGATATGAC 7428
Db 7644 GCGCGAAGAAACAGAACTGCCCCATCAATGACTTAAGAACTGTGTGCTATGACACAC 7703
QY 7429 AACAGATGATCTGTCCACAAAGAGAGGCTCTTAAGGCTTAAAGGATTAATTTT 7488
Db 7704 AATTGTGTATTTCCACCACTCAGCAGTGTCTTGCAGAAAGCAGAAAGTACATTT 7763
QY 7489 GATAGATGCAAGTCTCGACTTCTTACAGACTCAGTCTTAAAGGACTTAAGCTAGCG 7548
Db 7764 GACAGACTGCAAGTCTTGGACAGCCATTAACAGAGCTACTCAAGAGGTTAAAGACGG 7823
QY 7549 GCTTCCAAAGTCAACCCAGGCTCTCACCATGAGAGAGGCTTGCAGTTAACCCCA 7608
Db 7824 GCGTCAAAAGTGAAGGCTAATCTTATCCGTAGAGAGAGCTTGCAGCTGACGCCCA 7883
QY 7609 CATTTGCAAGATCTAAATATAGGGTTTGGGGCTTAAGAGGTCCGACGTTGTCCGGAGG 7668
Db 7884 CACTCAGCCAAATCAAGTTTGTGTTATGGGGCAAAAGAGTCCGTTGCCATGCCAAGAG 7943
QY 7669 GCGGTTAACCATCAAGTCCGTGTGAAAGAGCTCTGAGAGACTCAGAAACACCAATT 7728
Db 7944 GCGGTAAACCATCAAGTCCGTGTGAAAGAGCTTCTGAAAGACATTAACCAATTA 8003
QY 7729 CCGAACAACATTATGGCCAAAAATAGAGTGTCTGTGTGACCCCAACAAGGGGGCAAG 7788
Db 8004 GACATCAACATCAATGCTTAAGAACGAGGTTTCTGTGTGACCTGAGAGAGGGGCTGCT 8063
QY 7789 AAGCAGCTGCGCTTATGCTTAACTCTGACCTCGGCTCAGGGTCTGCAGAAAGATGGCC 7848
Db 8064 AAGCAGCTGCTTCACTGTGTCTTCCCGATCTGGGCTGTGCGTGTGCGAAAAAGATGGCT 8123
QY 7849 CTTTATGACTTAACAAAAAATCTCTCAGCGGCTGATGGGGCTTCTTATGATTCAG 7908
Db 8124 TTGTACGACGTGTATCAAAAGCTCCCTTGGCCGTATGGGAAAGCTCTACGGAATTCAG 8183
QY 7909 TATTTCCCGCTCAGCGGGTGAAGTCTTCTTGAAGGATGGCGGAAAAAGAGACCT 7968
Db 8184 TACTCACAGAGACGGGGTGAATTCCTGTGCAAGCGTGAAGTCCAAAGAAACCCCA 8243
QY 7969 ATGGATTTTTCGTATATACCCGATGCTTGAATCAACCGTCACTGAGAGACATCAG 8028
Db 8244 ATGGGTTCTCGTATATACCCGCTCTTGAATCAACGACTCACTGAGAGGACATCCGT 8303
QY 8029 ACTGAGAGTCCATATATGCGGCTGTCTTGTGCCGAGAGGCCCACTGCCATACAC 8088
Db 8304 ACGAGAGAGCAATCTACCAATGTGTGACTGACCCCAAGCCGCTGCGCATCAAG 8363
QY 8089 TCGTAACTGAGAGACTTTTACGTGAGAGGCTTAATTTCAACAGCAAGGCGCAACCTGC 8148
Db 8364 TCCCTACCGAGAGGCTTTATGTGGGGCCCTCTTAACCAATTCAGAGGGGAGAACTGC 8423
QY 8149 GGGTACAGGCGTTGCGGCGCAGCGGGGTGCTACCACTAGCATGAGGAGAAACATCACA 8208
Db 8424 GGGTATGCGAGGTGCGGCGGAGGCGGAGCTGACAACTAGCTGTGTAAACCTCTCACT 8483
QY 8209 TGTACGTGAAGCTTAAAGCGCTTGTAAAGCTGCAAGGATATATGCGCCCAATGCTG 8268
Db 8484 TGTATATCAAAAGCCCGGAGGCTGTGAGGCGGAGGCTCAAGAGTCAACATGCTC 8543
QY 8269 GTATGCGGCAATCTTGTGTCTATCTCAAGAAAGCAGGGGACCAAGAGAGACGAGCGG 8328

Db	8544	GTGTGTGCGCAGCACTTATGTGCTTATCTGTGAAGCGCGGGGGTCCAGAGAGACGCGCG	8603
Qy	8329	AACTGAGAGGCTTTCACGGAGGCTATGACCAAGTATTTTGCCCTCTCTGTGTGACCCCCCC	8388
Db	8604	AGCTGAAAGGCTTTCACGGAGGCTATGACCAAGTATCTCGGCCCTCTGTGGGACCCCCCA	8663
Qy	8389	AGACCGGAGTATGATCTGTAGCTGTATCATCTTGCTCTCTCAAAATGTGTCTGTGGCGCTG	8448
Db	8664	CAACCAAGAAATACGACTTGGAGCTCATTAACATCATCTCTCTCCAGCTGTCACTGGCCAC	8722
Qy	8449	GGCCCAACAGGCGCGCCGACGATACTACCTGACACAGAGCCCTACACTCCAAATGCGCCGG	8508
Db	8724	GACGCGCTGTGAAGAAGGGTCTACTACTCACTCAACCGGTGACCTTACAACCCCTCGCAGAG	8788
Qy	8509	GCTGCTCGGGGAAACAGTTATGACATCTCCCTCTGCAATTTATGGCTGTGGAAACATATCCAG	8566
Db	8784	GCTCGTGTGGAGACAGCAAGACACATCCCAATCAATTTCTGTGGCTAAGGAACATATATCATG	8844
Qy	8569	TACGCCCCGACCATATGAGGCTCGCATGTGCTCTGATGACACACTTTCTTCCATTTCTCATG	8628
Db	8844	TTTGTCCCCCAACATCTGTGGCCAGAGATGATATCTGATGACCCATTTTCTTATAGGCTCTTATTA	8900
Qy	8629	GCTCAAGACACGCTGACACAGAACTCTCAACTTTTATGATGTATCGAGACCGGTACTCCGCTG	8688
Db	8904	GCCAGGGACACAGCTTGAACAGGCGCTCATTTGCGAGATCTAGCGGGCTGTGTACTCATTA	8963
Qy	8689	AGTCCCTTGGACCTTCCAGCTTATATGAAAGGTATACATGGGCTTGAACGCTTTTCTCTG	8748
Db	8964	GAACCACTGTGATCTCACTCCATCAATCATTCMAAAGACTCCATGCGCTCAAGCGCATTTTCACTTC	9022
Qy	8749	CACACATATCACTCCCAACGAACTGACAGGGGTGCTTACGCGCTCTAGAAAACCTTGGGGCG	8808
Db	9024	CACAGTATCTTCCAGGTGAATTAATATGGGTGGCCGATATGCTCAGAAAACCTTGGGGTAT	9083
Qy	8809	CCACCCCTCAGAGCGGTGGAAGACCGCGGACATGTGACATCAGAGCGCTCCCTCATCTCCGCT	8868
Db	9084	CCGCGCTTTCAGAGCTTGGTGGACACCGGGCCGAGACGTCGCGCTTCTGGCCAGA	9143
Qy	8869	GGGGGAGAGGGGCGCTTGTGCGTTCGATATCTTCTTATTTGGGCGCGTGAACAACAAAGTTC	8928
Db	9144	GGAGGAGGGGTGCGCATATGTGGCAAGTACCTTTCMACTGGGCACTAAGAACAAAGCTTC	9203
Qy	8929	AAACTCACTCATTTGCGGAGACGCGCTCTCGATTTATATCAGCTGTGTTACACGTCGCGC	8988
Db	9204	AAACTCACTCCAAATAGCGGCCGCTGTGGCCAGCTGAACTTGTCCGGCTGTGTTTACACGCTGGC	9266
Qy	8989	GCCGCGCGGGGCGCAATTTATCAACGCGTGTGCGGTGCGGACCCCGCTTATTTGCTCTTT	9048
Db	9264	TACAGCGGGGAGACATTTATCAACGCGTGTCTCATGTCCGGGCCCGCTGTGATCTGGTTT	9322
Qy	9049	GGCTACTCTCTATTTTGTATGGGGGTAGGCTTTTCTAATCTCCCGCTCG	9098
Db	9324	TGCCATCTCTGTCTTGCGACGGGGTAAAGCACTAATCTCTCCCAACCG	9373

```

RESULT 13
US-09-388-874-1
: Sequence 1, Application US/09388874
: Patent No. 6284249
: GENERAL INFORMATION:
: APPLICANT: Veronique Barban
: TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR
: FILE OF INVENTION: TREATING C HEPATITIS
: FILE REFERENCE: PWC97-03A
: CURRENT APPLICATION NUMBER: US/09/388,874
: CURRENT FILING DATE: 1999-09-02
: EARLIER APPLICATION NUMBER: PCT/FR98/00448
: EARLIER FILING DATE: 1998-03-06
: EARLIER APPLICATION NUMBER: 97/02,887
: EARLIER FILING DATE: 1997-03-06
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1

```

```

; LENGTH: 9379
; TYPE: DNA
; ORGANISM: Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (320) ... (9352)
US-09-388-874-1

```

Query Match	45.2%	Score 4115.2	DB 3	Length 9379
Best Local Similarity	66.3%	Pred. No. 0		
Matches 6042; Conservative	0	Mismatches 2978	Indels 90	Gaps 5

OY	1	ATGAGCAAAATCTCTAAACCTCAAGAAAAACCAAAAGAAACACCAACCGTGGCCCAAA	60
Db	320	ATGAGCAAGAAATCTTAAACCTCAAAAAAATAAAGCTAACCAACCGTGGCCCAAG	379
OY	61	GACGTTAAGTTTCCGGGCGGCGGCAGATCGTTGGCGGATATACCTTGTGCGCGCAG	120
Db	380	GACGTTCAAAGTTCCCGGGGCGCGGTCAAGTCTGTGTGGAAGTTTACTTGTTCGCGCAG	439
OY	121	GGCCCCAAGTTGGGTGTGCGCGCAACAAGAAAGCTTGGAGCGGTCCAGCCACGTGGA	180
Db	440	GGCCCTAGATTGGGTGTGCGCGCAACAAGAAAGCTTCCGAGCGGTGGCAACTCGAGGT	499
OY	181	AGGGGCCAGCCATCTCTTAAAGATCGGCGCTCACTGGCAAAATCTGGGGAAAAACAGA	240
Db	500	AGAGGTCAAGCTTATCCCAAGGCTGTGTGGCCCAAGGACAGAACTTGAGCTCAACCCGG	559
OY	241	TACCCCTGGCCCTATACGGGAATGAGGAAATCGGCTGGGAGAGATGGCTCTGTGCCCC	300
Db	560	TACCTTGGCCCTCTATGGCAATGAGGAGCTGGGGTGGCGGAGATGCTCTGTCTCC	619
OY	301	CGAGTTTCCGCTCCCTCTTGGGGGCCCAATGACCCCGGCATAGGTGCGCAAGTGGT	360
Db	620	CGTGGCTTCGGCTAGCTGGGGGCCCAAGAACCCCGGCGTATGTGCGCAATTTGGGT	679
OY	361	AAGGTATCGATACCTTAAAGTGGGCTTTGGCACTCATGGGGTACATCCCTGTGTG	420
Db	680	AAGGTATCGATACCTTAAAGTGGGCTTTGGCCACACTCATGGGGGTACTACCGTCTGT	739
OY	421	GGGCCCCGCTCGGCGGGGTGCGCAGAGCTTGGCGCATGGCGTGGAGAGTCTGGAGAC	480
Db	740	GGGCCCCCTCTTGAAGGGGTGCGCAAGGACCTTGGCGATGGCGTCCGGTTCTGGAGAC	799
OY	481	GGGTTATATTTGGCAAGAGGAATTAACCCGGTGTCTCTTTTCTATCTTCTTGCGCC	540
Db	800	GGCGTGAATATGCAACAGGAACTTTCGTGGTGTCTTCTCTATCTTCTTCTGCGC	859
OY	541	CTGTGTCTCGATACCAACCCCGGCTCTCGGCTGCGGAATGAAAGAACATCAGTACCGGC	600
Db	860	CTGTCTCTTGTGCTTGACTGTGCGCGTTCGCGCTACCAAGGCGCAACTCCACCGGCTT	919
OY	601	TACATGTGTACTACACTGACCAATGACAGATTAACGTGGCAGCTCCAGGCTGTCTG	660
Db	920	TACCACTGACCAATGATGTGCTTAACCTCGATATTTGTATCGAGCGGCGCGATGCATC	979
OY	661	CTTCACGTCGCCGGGTGCTGCCCTGTGCGAAGAAATGGGAAATGATCTCACTGTCTGATA	720
Db	980	CTGACACTCCGGGGTGTGCTCTTGGCTTCGTATAGGGAACGCTCGAGGTGTGGGTG	1039
OY	721	CCGGTCTCACCGAATGTGGCCGTGACGCGCGCGCCCTTCACGCAAGGGCTTGGCGAGC	780
Db	1040	GCATATGACCCCTTACGGTGGCCACCAAGGAATGGCAAACTCCCGCGACGCACTTCAAGT	1099
OY	781	CACATTCGACATGTTGTGTATGTCCGCGACAGCTCTGTCTGTGCCCTCTAAGTGGGGAACTC	840
Db	1100	CACATTCATCTGCTGTGTGCGGAGGCGCAACCTCTGTTTGGCCCTCTAAGTGGGGAACCTA	1159
OY	841	TGCGGTGGGGTGATGCTCGACGCCCAATATGTTCAATTGCTCGCGCGACACATCTGTTTT	900
Db	1160	TGCGGGTCTGTCTTCTTGTGTGGCAACTGTTCACCTTCTCTCCAGCGGCACTGAGCG	1219
OY	901	GTCCAGACATGCAATGTCTCATCTTACCTTGGTACCATCACTGAGACACCGATGCAATGG	960

1220 ACCGAGGTGCAATGCTCTATCTATCCCGGCATATAACGGGATACCGCATGGGATG 1279
QY GACATATGATGAATGCTGCGCCACGGGCTACCATATCTTGGCGTACGGATGCGTGC 1020
Db GATATGATGATGAATGCTGCGCCACGGGCTGATATGCTGCTGCGGATC 1339
QY CCCGAGTCAATTAAGACATCATTAAGCGGGCTCATTTGGGGCGTCAATGTTCCGCTGGCC 1080
Db CCACAGCCCATCTTGGACATGATCGCTGATGCTCATCGGGAGTCCCGGGCGCATAGGG 1399
QY TACTTCTATGACGAGGAGCGTGGCGGAAGTGTGTCTCTTCTGTTGGCCCGGG 1140
Db TATTTCTCCATGCTGGGGAATCGGCGCAAGTCTGTGATGCTGCTGCTATTTGGCCGG 1459
QY GTGAGCGCGCGCAACCCATCTGTTGGGGGTTCTGGCGCGGACACACCGGGGCGCTCAC 1200
Db GTGAGCGCGGAACCCACGTCACCGGGGGAATGCCGCCACCTGTGTCTGATTTGTT 1519
QY AGCTTATTTGACATGGGCGCCGACGAGAAATCCAGTCTGTTAACCAATGCGAGCTGG 1260
Db AGCTTCTGACACAGGCGCCGACGAGAAATCCAGTCTGATTAACCAACGCGAGTTGG 1579
QY GACATCAACCGCACCGGCTGAACTGCAATGATCTCTTGGACACCGGCTTTATCGGCTC 1320
Db CACCTCAATGACACCGGCTGAACTGCAATGATGCTCTCAACACCGGCTTGGGACGG 1639
QY CTGTTCTACACCCACAGCTTCAACTGCTGAGGATGTCGCGAGGCAATGTCGCGCTGG 1380
Db CTGTTCTATACCAACAGTTCAACTCTTCAGGCTGTCTGAGAGGCTAGCGAGCTGCGA 1699
QY AGTATCGAGGCTTCCGGATGGGATGGGGCGCTTGCAATATGAGATTAATGTACCAAT 1440
Db CCCCTTACCGATTTTGAACAGGGGCTGGGCGCTTATGATGCAACGGAACGGGCGC 1757
QY CCAGAGATATGAGACCTCTATTTGCTGGCACTACCCACCAAGGAGTGGCGGTGCTC 1500
Db CCAGACGAGGCGCTTACTGCTGACCTACCCCCCAAAACCTTGGCGTATTTGGCC 1813
QY GCGAAGCTGTGTGGCGCCAGTGTACTGTTTCAACCCCGACGAGTGGATGGGACG 1560
Db GCGAAGTGTGTGTGGCGGTATGTTCTTCACTCCAGCCCGTGGTGGGACG 1873
QY ACCGACAGGCTTGGAGCGCCACTTACACGTGGGGGGAATGACAGATGTCTTCTTA 1620
Db ACCGACAGGTCCGGCGCGCCACTTACAGCTGGGGGTAAATGATACGACGTCTTCGTC 1933
QY TTGAACAGCACTGACCAACCGCTGGGATCATGTTGGCTGACAGTGGATGAATCTTCT 1680
Db CTTAACAAATACAGGCGACCGCTGGGCAATTTGGTTCGGTTGTACTGGATGAATCAACT 1993
QY GCGTACACCAAGACTTGGGGGACCAACCGTGGCGGTCTAGAGTGACTTCAACGCGACG 1740
Db GGAATTCACCAAGTGTGGGAGCGCTTCTTGTGTACTGGAGGGCGGGCAACACA 2051
QY ACGGACTGTGTGGCGCCACGACGACTGTTTAAAGAGCATCTGTATACACTTACCTCAA 1800
Db CCTTGCACGTGCCCACTGATGTCTCCGGAAGCATCCGACGCGCATATCTCCGG 2107
QY TGGGCTGTGGGCGCTTGGCTTCAAGCCAAAGTGTGTATGATACCTCTACAGCTCTGG 1860
Db TGGGCTGTGGGCTTGGCTTCAACCCAAAGTGTGTATGATACCTCTACAGCTCTGG 2167
QY CATTACCCCTGACAGTTAATTAATACCATCTTCAAAATAAGATGATGTTGGAGGGGTT 1920
Db CATTATCTTGTATCACTCACTACATCAATATTTAAATCAAGATGATCGTGGAGGGGTC 2227
QY GAGACAGGCTCAACGCGTGAATGCAATTTTCACTGTGGGATGCTGTGCAACTTGGAGAC 1980
Db GAGACAGGCTGAGAGGCTGTGCACTGAGCGCGGGCGGAAGTTGCGATCTGGAGAGAC 2287
QY AGAGACGAAGTCAACTGTCTCTTGTGTCATCCACACGGAATGGGCAATTTTACCT 2040
Db

2288 AGGACAGGTCCGAGCTACGCCGTTACTGTACCACTACACAGTGGAGGTCTCCCG 2347
QY TGCCTTACTCGGACTGTGGCGCCTTGTGTGACTGTGCTTCTCCACTCCACCAAAATC 2100
Db TGTCTTCTACCAACCTTACAGCTTGTGTCCAGGCTCATCTCCTCCACGAAACAT 2407
QY GTGAGAGTACAAATGATATGAGGCTTATCAGCTGCCCTCAAAATATACGTCCGATG 2160
Db GTGAGAGTCACTATGTAACGGGGTGGGGGTCAACATGCTGCTGGGCGCATTAAGTG 2467
QY GAGTGGGTAAATCTCTTATCTGCTTTAAGGAGCGCCAGGGTTTGGCGCTTATAG 2220
Db GAGTGGGTGTTCTCTGTTCTTCTGCTTGAAGAGCGCGGCTGTGCTGTGTTGG 2527
QY ATGCTATCTTGTGGGCGAGGCGCAAGACGACTAGAGAACTGTGATCTTGTGACGCT 2280
Db ATGATGCTATCTATATCCCAACGGAGGCGCTTGGAGAACCTGTGATATCTTAATGA 2587
QY GCGAGCGAGCTAGCTGCAATGGCTTCTTAATTTGTGATCTTTTGTGAGCTGCTGG 2340
Db GCATCCCTGGCGGAGCGACGCTGTGATCTTCTGCTGTGCTTGTGCTTGGATGG 2647
QY TACATCAAGGTGCGGTAGTCCCTTACTATNTCCCTGACGCGCTGTGCTT 2400
Db TATTTGAAGGTAAAGTGGTGGCCGAGCGTGTACACTTCTACGAGATGAGCTTCTC 2707
QY AGCTTACTGCTCTAGCATTTGCCCAACAGGCTTATGCTTATGACGATCTGTGATGCG 2460
Db CTCTGCTCTGTGTGGCGTGTGCTCCAGCGGGGTAGCGGCTGAGACGGAAGTGGCGGG 2767
QY CAGATAGAGCGGCTGTGCTGTAAATGATCACTCTTTTACTCAACCCCGGATTAAG 2520
Db TCGTGGCGGTTGTGTTCTCGTGGGTTGATGGCGCTGACTGTGACATATTAAGA 2827
QY ACCCTTACGCGGTTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
Db CGCTATATAGCTGT 2887
QY GTCAAGAGTGGGACACACCTATGACAGTGGCGGCTGGCTGTAGTGGATCATATGGGCG 2640
Db GTGACAGTGGATTTCCCGCTCAACGTCGAGGGGGGCGGAGCGGCTATCTTAC 2947
QY GTGCCAAATTTTACCAGGT 2700
Db ATGT 3007
QY GGGCTGTACTCTCTTAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2760
Db GGAACCTTGTGATTTCTTAAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3067
QY GCTTACTAGAGATGTGACCAATGACAGGCACTTCCGCGGGGCGAGTACGTCCAGATG 2820
Db GGCCTTCTCCGTTCTGCGCGTGTAGCGGGAATGATGAGAGGCAATTAAGTGAAG 3127
QY GCGCTACTAGCCCTTGGCAGTGTGACTGGCACTTATATGACCACTTACCCCTATAG 2880
Db GTCATATTAATGTAGGGGCGCTTACTGACACTATGTTTATTAACCATCTCACTCTCT 3187
QY TCGGATTTGGGCTGTAGTGGCGCTTGGGGGCTGTGGGCGTGGCGGTGAGCTTATCTTC 2940
Db CGGAGCTGGGCGCAACGCGCTTGGGAGATGTGGCGGTGTAGGCAAGTGTCTTC 3247
QY AGTCCGATGAGAAAGAAATGCTATTTGTGGGAGCGGAGACAGCTGTGTGGGACAT 3000
Db TCCCAATATGAGACCAAGCTCATACGTGGGGGCGAGATACCGCGCGGTGGGATGAC 3307
QY TTACACGACTTCCGCTGTGCGCGGACTTGGTGGGAGGTCTCTTGGCCGACTGAT 3060
Db ATCAAGGCTTGTGCTTGTTCGCGCGGAGGGCGGGAATATCTGCTCGGCGCAAGCGAT 3367
QY GGTATATCTTCAAGGGGTGTGATCTTCTGCGCCCATCATGCTTAACGCCACGAGACA 3120
Db GGAATGTGTCCAAAGGGGTGTGAGTGTGTGGCGCCATCATCAGCGCTACCGCCAGACA 3427

QY 3121 CGTGGCCTTTTGGCACCATAGTGTGATGACGAGGCGCGACAAAGACAGAGCT 3180
DB 3428 AGGGGCTCTTAGGATGATATACCAAGCTTAAGCGGAGCAAAAACAAAGTGAAG 3487
QY 3181 GGGGAAATTCAGAGTCTGTCCACAGTCACTCAGTCTTCTCGAGACATCATCTCGAGG 3240
DB 3488 GGTAGAGTCCAGATGTGTGCACTGTGTGCCAAACCTTCTCGAGAGTGTGATCAATGGG 3547
QY 3241 GTTTTGTGACGTCTACCATGAGCTGGCAACAAGACTGTGGCCGGCTACAGGGGTCCG 3300
DB 3548 GTGTGTGAGACTGTCTACACAGGGGCGGAAACGAGACCAATCGGTCAACCAAGGTCTT 3607
QY 3301 GTACGAGATGTACTCAGTGTGAGGGGGGACTTATGAGGGTGGCCAGCCCCCTGGG 3360
DB 3608 GTCATTCAGATGTATACCAATGTATGACCAAGACCTTGTGGCTGGCCCGCTCGCAAGT 3667
QY 3361 ACTAAATCTTTGAGCCGTGACCGTGTGAGCGGTGACCTGTACCTGTGACCGGAGAC 3420
DB 3668 AGCCGCTCATTTGACACCTGTGACCTTGGGGCTCTCGGACCTTTACCTGTGTACGAGGAC 3727
QY 3421 GCTGATGTATCCCGGCTCGAAGACGCGGGGACAAAGGGAGCGCTACTCTCCGAGA 3480
DB 3728 GCGGATGTCAATCCCGGCGCGCGGGGTGATGACAGGGGACGCTGTGTGCGCCCGG 3787
QY 3481 CCTCTTTCACCTTGAAGGGGTCTCTAGAGAGCCCGGTGCTATGCCCCAGGGGCAAGCT 3540
DB 3788 CCATTTCTCACTTGAAGGCTCTCTCGGGGGGTCCGCTGTGTGTCCCGCGGGGACAGCC 3847
QY 3541 GTCGAGTCTTCCGGGAGCGTGTGTCTCTCGGGCGGTGAGCTAATGATTTGATC 3600
DB 3848 GTGGGCAATTTTGAAGGCGGGGTGTGACCCGTGTGAAGGTGAGTGAAGTATTATC 3907
QY 3601 CCGGTGAGACACTGCAATGTGTACGCGGTCCCCACCTTTAGTGAACAAGACACCA 3660
DB 3908 CCTGTGAGAACTTGAAGACCAACATAGGTCCCGGTGTTCACGAGTAATCTCTTCCA 3967
QY 3661 CCTGTGCCCCGAGACTATCAGGTGGGGTACTGTGATGCCGACCTGGCAGTGGAGAG 3720
DB 3968 CAGTATGTGCCCGAGACTTTCAGGTGTCTCACTCATGTCTCCACAGGAGGGGAGAA 4027
QY 3721 AGCACAAGTCTCTGTGATATGTCTCAGGGGTATTAAGTGTAGTCTTAATCCC 3780
DB 4028 AGCACAAGTCCCGGCTGATATGCACTCAGGGCTATTAAGTGTAGTCTTAATCCC 4087
QY 3781 TCAGTGTGACCACTCGGGGTTTGGGGCGTACTTGTCTAAGGACATGSCATCAATCCC 3840
DB 4088 TCTGTGTCTCAACACTGGGCTTGTGTCTTACATGTCCAAGGCTCATGGGATCGATCT 4147
QY 3841 AACATTTAGACTGAGTCAAGACTGTGACGACCGGGGGCGCCATCACTACTCCACATAT 3900
DB 4148 AACATTCAGAACCGGGGTGAGAACAATTAACAATGGAGCCCAATCAAGTACTCCACTAC 4207
QY 3901 GGCAAATTTCTCGCGAGTGGGGGCTGTGCGGGCGGCGCTACGACATCATCATATGTGAT 3960
DB 4208 GGCAGTCTCTTGGCGGACGGCGGGTCTCGGGGGGCGCTTATGACATTAATTTGTGAC 4267
QY 3961 GAATGCCATCGGTGACCTTACCAACATCTTGGGACATCGAAGACGCTTGTGATCAAGA 4020
DB 4268 GAGTGCACACTCAACGATGACCAATCATCTTGGGACATCGGCACTGTCTTGTGACAGCA 4327
QY 4021 GAGACGCTGGGGGACACTTAACGTGTGTGCTACACACTACGCCCCCTGGGTGAGTACA 4080
DB 4328 GAGATGTGGGGGAGAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4387
QY 4081 ACCCCCAACCCCAATAGAGAGTGGCCCTTGGGAGAGAGGCGAGATCCCCTTGTAT 4140
DB 4388 GTGCCCATCCCAACATCGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4447
QY 4141 GGGAGGGGCAATCCCTGTCTTACATCAAGGAGAGAAAGACATCTGTCTGTGATTTCA 4200
DB 4448 GGCAGAGCTATCCCTCGAAGTATCAAGGGGGGAGACATCTCATCTTGTGATTTCA 4507

QY 4201 AAAAAAGTGTGACGAGCTCGGCGCGCCCTTTCGGGGTATGGGCTTGAATCACTAGTGA 4260
DB 4508 AAGAGAGTGTGACGAGACTTCGCGCAAGCTGTGTGATTTGGGCTATTAATGCTGTGACC 4567
QY 4261 TACTAGAGGGGTTTGAAGCTCTCCGTAATACCACTCAGGAGAGAGTGTGTGTGTGCC 4320
DB 4568 TACTACGCGGCTTGTGACGCTGTCCGTATCCGACAGCGGGGAGTGTGTGTGTGTGTGCA 4627
QY 4321 ACCGACGCTCTATGACAGGATATCTGGGAGCTTTGATCTCCGATTCGATCTGACAGTA 4380
DB 4628 ACCGATGCTCTATGACAGGCTATACGAGGAGCTTGTGACCTGGGTGATGACTGCAATACG 4687
QY 4381 GCGTCACTCAAGTTTGAAGCTTCAAGTTTGAAGCCCAATTCACATTAACCAACAGATT 4440
DB 4688 TGTGTACCCAGACAGTGTGATTTCAAGCTTGAAGCTTCACTTCACTTGAAGACATACG 4747
QY 4441 GTCCCTCAGAGCGCTGTCTCAAGTACGACGCGCGGGGTGACAGGGTGAAGGAGACTG 4500
DB 4748 CTCGCCAGGATGTGTCTCCGCACTCAAGCTGTGGGGCAGGACTGGCAGGGGGAAGCCA 4807
QY 4501 GGCATTTATAGTATGTCTTCACTGTGAGCAGGCTCAGGAATGTTTGAACATGTAGTG 4560
DB 4808 GGCATCTACAGATTTGTGTGACCGGGGAGCGCCCTCGGCAATGTTTCACTGTCCGTCT 4867
QY 4561 CTCGTGAGTGTCTACGACGAGGGGCGGACTGTGTATGAGTCAACCATCGAGACACCC 4620
DB 4868 CTCTGTGATGTCTATGACGAGGCTGTGCTTGTGTATGAGCTTACGCGCCCGCAGACTTACA 4927
QY 4621 GTGAGGCTCAGGCGATTTTCAACAGCGCGGTTTGTGTGTGTGTGTGTGTGTGTGTGT 4680
DB 4928 GTTAGGCTACAGAGTATACATGAACACCCGGGGCTTCCGTGTGTGTGTGTGTGTGTGT 4987
QY 4681 TTTTGGAGGCAATTTTACCGGCTTCAACACATATATGATGCCACTTCTTCCCAACA 4740
DB 4988 TTTTGGAGGCGCTTCTTACAGGCTTCACTCAATATGATGCCACTTCTTATCCAGACA 5047
QY 4741 AAGCAATCGGGGGAATTTCCGATCTTAACAGGCTTACAGGCTTCAAGTGTGCGCTAGG 4800
DB 5048 AAGCAGATGGGAGAACTTCTTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5107
QY 4801 GCCAAAGCCCCCCCCGCTGT 4860
DB 5108 GCTCAAGCCCCCTCCCATGT 5167
QY 4861 ACACTGT 4920
DB 5168 ACCCTCAATGGGCAACACCCCTGTATACAGCTGTGGGCGTGTGTGTGTGTGTGTGTGT 5227
QY 4921 CTCACATCCGTTGAGAAATATCATCGCACCTGTGATGACGCGACTTGTGTGTGTGTGT 4980
DB 5228 CTGACGACCCAGTCAACCAATATCATATGACATGTGTGTGTGTGTGTGTGTGTGTGTGT 5287
QY 4981 ACCAGCAATGGGCTTGT 5040
DB 5288 AGGAGGACCTGT 5347
QY 5041 ACCGGGT 5100
DB 5348 ACAGGCTGT 5407
QY 5101 GACAGAGAGTCTCTATGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5160
DB 5408 GACAGGAGTCTCTTACGAGAGT 5467
QY 5161 CTCATTGAAGGGGAGCGGATGACCGAGATGCTGAAGTCCAAGTCCAAGGCTTATTTG 5220
DB 5468 TACATGAGAGAGGAGT 5527
QY 5221 CAGCAAGCTTCCAAAGAGTCAAGACATACAAACCACTGTGAGGCTTCAATGGCCCAAG 5280
DB 5528 CAGACGCGCTTCCGTAGGAGAGT 5587
QY 5281 GTAGAACAAATTTGGGCGCAACATGTGTGAACCTTATTAAGCGGCTCAATCACTGTGCA 5340

Db 5588 CTCAGACCTTCTGGGCGAAGCATATGTGAACTTCATCAGTGGGAATACATATTG3GC 5647
Oy 5341 GGACATCAACATGCGGAGGAGACCCCTGAGTAGATTCATGATGGGCTCAGGCGCC 5400
Db 5648 GGGTTGTAAAGCTGCTGTGTAACCCGCCATTTGCTTATGATGAGGCTTTTACAGCTGCT 5707
Oy 5401 CTCACAGTCCGCTGTCAACAAGACCACTATCCTTCTCAACATTTTGGGGGCTGAGCTA 5460
Db 5708 GTACACAGCCCACTAACCACTAGCCAAACCTCCCTTTCACATATTTGGGGGGTGGGCTG 5767
Oy 5461 GCATCCCAATTTGACACACCCCGGGGGCCACTGGCTTCCTTTGTCAGTGGCTTAGTGGGA 5520
Db 5768 GCTGCCAGCTGCGCGCCCGGCGCTCTAGCTGCTTTGTGGCGCTGGCTTGTAGCTGCGC 5827
Oy 5521 GCTGCCGTAGGAGATAGGCTTAGGTTAAGGTAGTGAAGATCCTGAGAGAGGCTTATGCT 5580
Db 5828 GCCGCATCGGAGGTGTTGAGCTGGGGAAGGTCTCTATAGACATCTTTGCGAGGTTATGCG 5887
Oy 5581 GCGGCGATTTGGGGGCTCTGCTGCGCATTTCAAGATCATGTGCGAGAAAGCCCTCGATG 5640
Db 5888 GCGGCGTGTGGGAGGCTCTTGCGCATTTCAAGATCATGAGCGGTGAGGTGCCCTCCACG 5947
Oy 5641 GAGATGTCTCAACTTCTGCTGCGAATTTCTGCTCGGGTGCCTTGTGTAGTGGAGTCTC 5700
Db 5948 GAGGACCTGTCAATCTACTGCGCGCATCTCTGCGCGGAGCCCTGTAGTGGGCTG 6007
Oy 5701 ATCTGCGGCGCATTTCTGCGCGAGACAGTGGAGACCGGGGGAAGGCGCGTCCATTTGATG 5760
Db 6008 GTCTGTGAGCAATATCTGCGCGGACGCTGTGCCCGAGGAGGAGGAGCTCACTGATGGATG 6067
Oy 5761 AATAGACTCATTTGCTTCTTCTTCCAGAGAAATCAAGTGCGCCCCACCACTAGTGAAG 5820
Db 6068 AACCGGCTGATAGCTTCTGCTCTCCGCGGGAACCATGTTTCCCCAGCACTAGTGCAG 6127
Oy 5821 GAGTGGATGCGTGGCAGCGGTGTACCCAACTATTTGCTCTTCAATTAACAGCTG 5880
Db 6128 GAGAGCGATGACGTGCGCGGCTCACTGCACTCAAGACAGCTCACTGTAACCAAGCTC 6187
Oy 5881 CTCAAGAACTCCAAACATGATTTCTGAGGACTGCGCCCATCCATCGAGGCGGCTGCG 5940
Db 6188 CTGAGGCACTGACCAAGTGAATAGCTCGAGGTGTACCACTCCATGCTCGGTTCTTGG 6247
Oy 5941 CTCGCGATGTGTGGAGCTGGGTTTGCACATCTCAACAGACTTTAAATTTGAGTGAAG 6000
Db 6248 CTAAGGGAATCTGGGAGCTGGATGATGAGGTTGAGCACTTTAAGACTGCTGCTAAA 6307
Oy 6001 TCCAAATTTATCCCAAAGATGCGCGCTCTCTTGTCTCTGCTCAAAAGGGGTACAAG 6060
Db 6308 GCTAAGCTCATGCGCACAGCTGCTGGGATCCCTTTGTGCTGCTGCGAGCGCGGTATAG 6367
Oy 6061 GCGCGTGGGCGGAGCTGGGATCATGACACAGGAGTCCCTTGGGGGCAATATCTCT 6120
Db 6368 GGGGTCTGCGAGTGGAGCGGATCATGACACTGCTGCCACTGTGGAGCTGAGATCACT 6427
Oy 6121 GGCATATGTCGCTTGGGCTCCATGAGATCAAGGGGCTTAAGACTGATGAATATCTG 6180
Db 6428 GGAATGTCAAAAAGGAGCGATGAGATCGTGGTCTTAGAGCTTGAAGAAATGTGG 6487
Oy 6181 CAGGGGACCTTTCTATCAATTTGTTACAGGAGGGCCAGTGCCTGCGAAACCGCGCA 6240
Db 6488 AGTGGGACCTTCCCATTAATGCTTACACAGGCGCCCTGTACCCCTTCTGCGCGG 6547
Oy 6241 AACTTAAAGTGGCATGAGGAGGTGGGCGCTCAGAGTACGGGAGTGAAGCGAGAC 6300
Db 6548 AACTTACAGTTGCGGCTATGAGGAGTGTCTGAGAGAAATATGTGAATTAAGGAGGTG 6607
Oy 6301 GGGTCATACCACTAATAACAGAGTCACTGATTAATTTGAAAGTCCCTGCGCACTA 6360
Db 6608 GGGGACTTCACTAAGTGAAGGTTAGTACTGACATCTCAATATGCGCGTGCAGAGTCTC 6667
Oy 6361 CCTTCTCCGAGTTCTTTTCTGGGTGAGCGAGTGCATAGTTTGCCTCCCA 6420

Db 6668 CCATGCGCGAAATTTTTCACAGATTTGACGAGGTGCGCTCACTAGATGTTTGGCCCCC 6727
Oy 6421 CCGAAGCCGTTTTTCCGGAGTGAAGTCTGTTCTGCTTGGGCTTAATCATTTGTCGTC 6480
Db 6728 TGCAGGCTTGTGTGGGAGAGATATCATTTAGATAGAGTCCACGAATACCGGTA 6787
Oy 6481 GGGTCCAGCTTCTTGGACCTTGAACCTGACACAGAGTATTTGATGCTATAC 6540
Db 6788 GGGTCCGAATTAACCTTGGACCGGACCGAGCGTGGCGTGTGACGTGCTCACT 6847
Oy 6541 GATCCATCTCATATCAAGCGGAGAGTGAAGCGCGGCTTGAACCGGGGGGTCAACCCCA 6600
Db 6848 GATCCCTCCATATAACAGACAGGCGCGCGGAGGTTGGCGAGGAGATACCCCC 6907
Oy 6601 TCCGAGGCAAGCTCCTGCGAGCGAGCTATCGGCACATCTGCTGAGACCACTGAC 6660
Db 6908 TCTGTGGCAGCTCCTCGGTAGCCAGCTATCCGCTCATCTCTCAAGGCAACTTGCAC 6967
Oy 6661 ACCCAGGCAAGCCTATGATGTGAATGTGATGTCTAAGCT-----GTTT 6708
Db 6968 GCTAACATGACTCCCTGATGCTGAGCTCATGAGGCCAACCTCTATGAGGAGAGAG 7027
Oy 6709 ATGGGGGCGATGTGATCTGGATAGAGTGTGGGTCCAAAGTGTGCTTCTGATCTCTC 6768
Db 7028 ATGGGCGGCAACATCAACAGGTTGAGTCAAAAACAAAGTGTGATTTGATCTCTTC 7087
Oy 6769 GACCCAAATGATGCAAGAAAGAGCGAGCTTGAAGCTTGCATACATCAATATAGTCTC 6828
Db 7088 GATCCCTTGTGGCGGAGAGAGACAGCGGAGATTCGTTACCGGCAAAATCTCGCG 7147
Oy 6829 CCCAAGAGAGTTCCACAGCTTTACCGGCTGGGCAAGGCTGATTTACAAACCAAG 6888
Db 7148 AAGTCTCGAGATTCGCCAGGCGCTCGCTTGGGCGCGGAGCTATTAACCCCGG 7207
Oy 6889 CTGTGGAATCTGTAAGAAAGGCGAGATTAACAAACCGGCACTGTTGGGGCTGTGCTC 6948
Db 7208 CTAGTGAAGCTGTGAAAAAGCCGACTAGAAACACTGTGTGTCATGCTGTGCGCTT 7267
Oy 6949 CTTCTCTTGAAGAAACCCCGAGCGCTCCCGCAAGAGGCGCGGACAGTGGGCTTAAGT 7008
Db 7268 CCACTCTCAAAAGTCCCTCTGTGCTCTCGGCTCGGCAAAAGGAGTGTGCTCACT 7327
Oy 7009 GAGGACTCATAGAGATGCTCTTCAACAGCTGGCCATTAAGTCTTTGGCCAGCCCC 7068
Db 7328 GAATCAACCTATCTACTGCTTGGCGGAGCTGCCACGAAAGCTTTGGAGCTCTCA 7387
Oy 7069 CCAAGGGGAGTTCAGGCTTTTCCAGGGGGGCGGGCTGGCCGATTCGCGAGTCAAG 7128
Db 7388 ACTTCCGGCATTAAGGGGCAATACGACAACTCTCTGAGCCGCGC----- 7435
Oy 7129 CTTCTGATGATGTTGGCCCTTTTGGAGACAGGTTTCATCTTTTCAATGCCCCCTCGAG 7188
Db 7436 CTTTCTGGCTGCCCCCGCACTCGAGCTGAGTCTTATTTCTTCCATGCCCCCTCGAG 7495
Oy 7189 GGGGACTTGGAGATCAAGACTGAGCTGAGCAGTGAAGAGCCCAACCCCCCCAG 7248
Db 7496 GGGGAGCTGGGAGTCCGA----- 7515
Oy 7249 GGGGGGGTGGAGCTCCCGGCTCGAGCTGGGGTCTGTGTTACTTGTCTCGAGAGAGAC 7308
Db 7516 -----TTTAAAGCGGGTCAATGTCTAACGCTCAGTGAAGGCAACGCG 7561
Oy 7309 GACTCGGTGATGCTGCTCATATGATCTCTGAGACGGGAGCTTAATTAATCTCTGT 7368
Db 7562 GAGGATGTGTTGTGCTGTCAATGTCTTACTCTTGAACAGGCGCACTGTCACCCGTG 7621
Oy 7369 AGTCCGAAGAGGAAGTTACCGATTAACCCCTTGAAGCACTCCCTTGTGCGATATCA 7428
Db 7622 GCGCGGAAGAACAGAACTCCCATCAATGCACTAAGCAACTGTTGCTACGTCACAC 7681
Oy 7429 AACAAAGTGTACTGTACCAACAAAGAGCGCTTCACTAAGGCTTAAAGGTAATTTT 7488
Db 7682 AATTTGTGTATTCACACACTTACGAGGTGCTTGCACAAAGGAGAAAGTCACTTT 7741

OY	7489	GATAGAGATGCAAGTGCAGCTCCCTACTACAGCTCAGCTTAAAGAAATTAATTAAGTACG	7548
Db	7742	GACGACCTGCAAGTTCTTGACAGCCATTACAGAGACGTACTCAGAGAGTTTAAAGCACG	7801
OY	7549	GCCTCGAAGGTACCCGCAAGGCTCCTCCACATGAGAGAGGCTTGCCAGTTAACTCCACCC	7608
Db	7802	GCCTCAAAAGTGAAGGCTTAACCTTGCTACTCGTAGAGAAAGCTTGACGCTCAGCCGCCCA	7861
OY	7609	CATTCTGCAACATCTAAATATATGGCTTTGGGGCTAAGAGGTCCGACGTTGTCCGGAGG	7668
Db	7862	CACTGAGCCAATCCCAAGCTTTGGTTATGGGGCAAAAGCGTCCGTGTGCATGCCAGAAAG	7921
OY	7669	GCCGTTAAACAATCAAGTCCGTGTGGAAGACCTCCTGAGAGSACTCAGAAACCAATT	7728
Db	7922	GCCGTAAACCAATCACTCCGTGTGGAAGAACCTTCTTGGAAGCAATGTAAACACATA	7981
OY	7729	CCCAACACATTATATGGCCAAAATGAGGTGTTCTGCTGGAACCCCAACAGGGGGCAAG	7788
Db	7982	GACACTACATCATGTGCTTAAGAACGAGGTATTTCTGCTGTACGCTGAGAAAGGGGGTCTG	8041
OY	7789	AAAGCAGCTCGCCTTATGTTTACCTGACCTCCGCGCTCAGAGTCTGGAAGATGAGCC	7848
Db	8042	AAAGCAGCTCGCTCATCTGTTTCCCGCATCTGGCGCTGCGCGTGTGCGAAAAGTGGCT	8101
OY	7849	CTTATATGACATTAACAACAAAATTCTCCAGGCGGTATGAGGGGCTTCTTATGATTCAG	7908
Db	8102	TTGTACGACGTGTTAACAAAGCTCCCTTGCCCGGTATGGGAAGTCTCTACGATTCMA	8161
OY	7909	TATTTCCCGCGTCAAGCGGGTGAAGTTCCTTTGAAAGCATGGGGGAAABAGAAGCCT	7968
Db	8162	TACTACCAAGAACAGCGGGTGAATTCCTGTGCAAGGTGGAAGTCCAAAGAAACCCCA	8221
OY	7969	ATGGGTTTTGTATGATATACCGAATGCTTTGACTCAACCTGCATGAGAGAGACATCAG	8028
Db	8222	ATGGGTTCTGTATGATATACCGCTGCTTTGACTCACAAGTCACTGAGAGGAGCATCCGT	8281
OY	8029	ACTAGAGATCCATATATCGGAGCTGCTCTTGCCCGAGAGGCCCACTGCCATACAC	8088
Db	8282	ACGAGAGGAGCAATCTACCAATGTTGTGACCTCGACCCCAAGGCCGCGTGGCATCAG	8341
OY	8089	TGCGTAATGAGAGACTTAAGTGGGAAGGCGCTAATGTTCAACAGAGGGCAAACTCG	8148
Db	8342	TCCCTCAACGAGAGGCTTATATGTGGGGGCGCTTACCAATTCAGAGGGGGAAGATCG	8401
OY	8149	GAGTACAGGCTTGCGCGCCAGCGGGGTGCTCAACTAGCATATGGGAAACCATCACA	8208
Db	8402	GGCTATGCAAGTGTGCGCGAGGGGCGCTATGACAACTAAGCTGTGTGAACCTCTACT	8461
OY	8209	TGCTAAGTGAAGCTTATAGCGGCTTTGTAAAGCTGACGGGATTAATGCGCCCAATGTG	8268
Db	8462	TGCTACATCAAGGCGCGGCGAGCGTGTGACCGGAGGCTTCAGGACTGACCATGTCTC	8521
OY	8269	GTATGCGCGATGACTTGTTGTCACTCAAGAAACCAAGGGGACCGAGAGAGACGAGCG	8328
Db	8522	GTGTGTGCGAGACTTATAGTGTGTTATCTGTGAAGCGGGGGGTTCAGAGAGACGGCG	8581
OY	8329	AACTGAGAGCTTTCACGAGGCTATGACACAGTATTTGCCCCCTCTGTGTACCCCCC	8388
Db	8582	AGCTGAGAGCTTTCACGAGGCTATGACACAGTATTTGCCCCCTCTGTGTACCCCCC	8641
OY	8389	AGACCGAGTATGATCTGAGCGTGATTAACCTTGCTCTCAAAATGTCTGTGGCGGT	8448
Db	8642	CAACCGAATACGACTTGAAGTCAATACATATATCTCTCAACGTATCAGTGGCCAC	8701
OY	8449	GGCCCAACAGGCGCGCAGATCTTACCTGACACAGAGCCCTACACTCAATGCCCCG	8508
Db	8702	GACGGCGCTGGAAGAGGGGTCTACTTACCTCACCCGTGACCTTACAAACCCCTCGCAG	8761
OY	8509	GCTGCGTGGGAAACGTTTACATCTCCCTGTCAATTAATGGCTGGGAAACATCATCCAG	8568
Db	8762	GCTGTGTGGAGACGACGAAAGACATCTCAAGTCAATTTCTGTGTAGGCAACATATATCAT	8821

Oy	8569	TACCCCCGACATATGGGCTCGAGATGGTCTGTATGTACACACTTCTTCTCATTTCCATG	8628
Db	8822	TTTGTCCCCCACAACGTGTGGGCGAGGATGATATCTAGTACGCCATTTCTTTAGCGTCTTATA	8881
Oy	8629	GCTCAGAACACGCTGTGACCCAGAACCTTCACTTTGATGATGTAGGAGCGGCTGTACTCCGCTG	8688
Db	8882	GCCAGGGACCGAGCTGTGAACAGGGCCCTCGATTGGCCAGATCTACAGGGGCGCTGTACTCCATA	8941
Oy	8689	AGTACCTTTGACCTCCAGCTATATATTTAAGTTACATGTGGGCTTTGACGCTTTTCTCTG	8748
Db	8942	GAACTACTGTATCTTACTCTCCATATCTTCATAGATCTCCATGGGCTTCAGCGCATTTTCACTC	9001
Oy	8749	CACACATTAACACTCCCAACGAACTGACACAGGCTGTCTTCAAGCTTCAGAAAATTGGGCG	8808
Db	9002	CACAGTTACTCTCCAGGTGAATTTAATTAAGGCTGTGCCGATGCCATGCCACAGAAAATTGGGCGTA	9061
Oy	8809	CCACCCCTCAGAGGGTGGAAAGAGCGCGGGCACGTCAGTCAGGGGGCTCCCTCATCTCCCGCT	8866
Db	9062	CCGCCCTTTGGCAGCTTTGGAGACACCGGGCCCGAGCGCTCCGCGCTTAGGCTTCTGGCCAGA	9121
Oy	8869	GGGGGAGAGCGCGCGTTTGGCGTGCATATCTTCTTCAATTTGGGCGGTGAGAACCAACGCTC	8928
Db	9122	GGAGGCGAGGGCTGCCATATGTGGCAAGTAACTCTTCTTCAACTGGGCGATGAAACAAAGCTC	9181
Oy	8929	AAACTCACTCATTTGCCGGGAAGCGCGCTCTCGATTTATCACTGTGTTACAGCTGGCG	8988
Db	9182	AAACTCACTCCAAATAGCGGCGCTGGCCAGCTGGACTTTGTCCGGCTGGTTTACGGCTGGCC	9241
Oy	8989	GCCGGCGGGGCGACATTTATCACAGCGTGTGCGGTCGCCGACCCCGCTTATTGCTCTTT	9048
Db	9242	TACAGCGGGGAGACATTTATCAAGCGTGTCTCATGCCCGGCCCGCTGGATCTGATTT	9301
Oy	9049	GGCTTACTCTTACTTTTGTATGGGGTATAGGCTTTTCTTACTCCCGCGCTCG	9098
Db	9302	TGCCTACTCTGTCTTGCGACGGGGTATGGGATCTTACTCTTCCGCCAACCG	9351

```

RESULT 4
US-09-916-359-1
; Sequence 1, Application US/09916359
; Patent No. 6538123
; GENERAL INFORMATION:
; APPLICANT: Veronique Barban
; TITLE OF INVENTION: VACCINE COMPOSITION F
; TITLE OF INVENTION: TREATING C HEPATITIS
; FILE REFERENCE: PMCE97-03A
; CURRENT APPLICATION NUMBER: US/09/916,359
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 09/388,874
; PRIOR FILING DATE: 1999-09-02
; PRIOR APPLICATION NUMBER: 97/02,887
; PRIOR FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 9379
; TYPE: DNA
; ORGANISM: Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (320) ... (9352)
US-09-916-359-1

```

Query Match	45.28;	Score 4115.2;	DB 4;	Length 9379;
-------------	--------	---------------	-------	--------------

Matches 6042; Conservative 0; Mismatches 2978; Indels 90; Gaps 5

1 ATGAGCACAATCTAACCCTCAAGAAGAAACCAAGAACAACCAACCGTGCACCAA 60

Db 320 ATGAGCAGCAATCCTAAACCTCAAAAAAAAACAACGTAACACCAACCGTCGCCACAG 379

61 GACGTTAGTTTCCGGCGCGCCAGATCGTTGGCGAGTATACTTGTGCCGCGCAGG 120

Db 380 GACGTCAAGTTCCCGGGTGGCGGTCAAGTCGTGGTAGATTCTTGTTCGCCGAGG 439
Oy 121 GAGCCGAGTTGGGTGTGCGCGCAAGAGAACTTCGAGCGGTCCAGCGAGTGA 180
Db 440 GGGCTAGATTGGGTGGCGCGCAAGAGAAAGACTTCCAGCGGTCCAGACTCGAGGT 499
Oy 181 AGGCGCAGCCATCCCTAAAGATCGCGCTCACTGGCAAACTCTGGGAAAAAGAGA 240
Db 500 AGAGCTCAGCTATCCCAAGGCTGTGGCCCGAGGGAGAGACTGGGCTCAGCCGGG 559
Oy 241 TAACCCCTGGCCCTATAGAGGAATAGAGGACTCGGCTGGGAGAGATGGCTCTCCCC 300
Db 560 TACCCTTGGCCCTATAGAGGAATAGAGGACTCGGCTGGGAGAGATGGCTCTCTCC 619
Oy 301 CGAGGTTCCGCTCCTTCTTGGGGCCCAATGACCCCGGCAATAGTCCGCAAGTGGT 360
Db 620 CGTGGCTCTGGGCTAGGTCGGGCGCCACAGACCCCGGCGTAGGTCCGCAATTTGGT 679
Oy 361 AAGGTCAATGATACCTTAACGTGCGGCTTTGGCCGACTCATGGGGTACATCCCTGTGTG 420
Db 680 AAGGTCAATGATACCTTAACGTGCGGCTTTGGCCGACTCATGGGGTACATCCCTGTG 739
Oy 421 GGGCGCCGCTCGGGGGGTGGCGCAGAGCTCGCGGCAATGGCGAGAGTCTGAGAGAC 480
Db 740 GGGCGCCCTCTTGGAGGCGTGCAGAGGCCCTGGCGCAATGGCGGTCTGAAAGAC 799
Oy 481 GGGGTTAATTTTGCAGACAGGAACTTACCAGGTGCTCTTTCTATCTTCTTGGCC 540
Db 800 GGGGTTAATTTTGCAGACAGGAACTTACCAGGTGCTCTTTCTATCTTCTTGGCC 859
Oy 541 CTGCTGCTCTGATCAACACCCGGTCTTCGCTGCGAATGAGAAATCATGATCGGCG 600
Db 860 CTGCTGCTCTGATCACTGTCGCGCTTCGCGCTCAAGATGCGCAATCCACGCGGCTT 919
Oy 601 TACATGATGATCAAGACCTGACCAATGACAGCATTAATGAGAGCTCCAGGCTCTGTC 660
Db 920 TACACGATCAACCAATGATTTGCTTAATCTGAGTATTTGTAAGAGCGCGCGATGCTTC 979
Oy 661 CTCCAGTCCCGGGTGGCGTCCCGTGCAGAAAGTGGGAAATGCACTCACTGCTGATA 720
Db 980 CTGCACACTCCGGGGTGGCGTCCCTTGGCTGTGAGAGGCAACGCTCGAGGTGTTGGGTG 1039
Oy 721 CCGGTCTACCGAATGTGGCGGTGACGCGGCCGGCGCTTCAGCAGAGGCTTGGGAGC 780
Db 1040 GCGATGACCCCTTACCGGTGGCCACAGGATGGCAAACTCCCGCAGCAGCTTCGAGCT 1099
Oy 781 CAGATGACATGTTGATGTCGCGCAGCTCTGCTGCGCTTACGTTGGGGAGACCTC 840
Db 1100 CAGATGATCTGCTTGTGGGAGCGCACCTCTGTGGCCCTCTTACGTTGGGGAGACTTA 1159
Oy 841 TGGGTGGGGTGTGCTGCGAGCCCAATGTTCACTTGTCTGCGCAGCAGCAGCTGTTT 900
Db 1160 TGGGGGTCTGTCTTCTTGTGGGCAACGTTCACCTTCTCTCCAGGCGGCACTGGAGC 1219
Oy 901 GTCCAAAGACTGCAATTTGCTCATCTAACCTGTACATCACTGACACCGCATGGCATGG 960
Db 1220 ACGCAAGGTTGCAATTTGCTCATCTAACCTGTACATCACTGACACCGCATGGCATGG 1279
Oy 961 GACATGATGATGAACGTGGCGCCAGGTCACATGATCTTGGGCTAGAGCATGGGTGTC 1020
Db 1280 GATATGATGATGAACGTGGCGCCAGGTCACATGATCTTGGGCTAGAGCATGGGTGTC 1339
Oy 1021 CCGGAGTCAATTAAGACATCAATTAAGCGGGCTCATTTGGGGCGTATGTTGGCTTGGCC 1080
Db 1340 CCACAGGCACTTGTGACATGATGCTGTCTCATCTGGGAGTCTTGGCGGCAATAGCG 1399
Oy 1081 TACTTCTCTATGACAGGAGCGTGGGCAAGTGTGTCATCTTCTGTGGCGCGCGG 1140
Db 1400 TATTTCTCATGATGTTGGGCAATGGGCGAAGGTCTGTGATGTCGTCTATTTGGCCGGC 1459
Oy 1141 GTGGAGCGCGCAGCATTAATGTTGGGGTCTTGGCGGAGAGACACCGGGGCGCTTACC 1200
Db 1460 GTGACGCGGAAACCACTGTCACCGGGGAGAGTGGCCGACACTGTGTCTGAGATTGTT 1519

Oy 1201 AGCTTATTTGACATGGGCCCCAGAGCAAAAATCAGCTGTTTAAACCAATGGCAGCTGG 1260
Db 1520 AGCTCTCTGACACAGGCGGCAAGAGAAAGTTCAGCTATATCAACCAAGGCAATGG 1579
Oy 1261 CACATCAACCGCAGCCGCTTGAATCTGCAATGACTCTTTGACACCGGCTTTTATGGTCT 1320
Db 1580 CACCTCAATAGCACCGGCTTGAATCTGCAATGAGCTCAACACCGGCTGTTGGAGAGG 1639
Oy 1321 CTGTTTCAACCCAGAGCTTCAACTGTCAGAGATGTCGCCAAGCAGTGTCCGCTGCGC 1380
Db 1640 CTTTTTCTATACCAAGATTTCAACTTTTCAAGCTGTCTGAGAGGCTAGCCAGCTGCCA 1699
Oy 1381 AGTATGAGGCTTCCGGGTGGATGGGCGCTTGGCAATATGAGATTAATGTACCAAT 1440
Db 1700 CCCCTTACGATTTTGAACAGGAGGCTGGGCGCTTATGATTAATGCAACGAAAGCGGCG-- 1757
Oy 1441 CAGAGATATAGAGACCTTATTTGCTGGACATACCCCAAGAGGAGTGTGGCGTCTGC 1500
Db 1758 ----CCGACCAAGGCGCTTATGCTGGACATACCCCAAAACCTTGGCGTATTTGGCC 1813
Oy 1501 GCGAAGACTGTGTGGCCAGTGTACTGTTTCAACCCCAAGCCAGTGGTATGGAGCAG 1560
Db 1814 GCGAAGATGTGTGTGTCCGGTATATTTGCTTCACTCCAGCCCGGTGTGTGGAGAG 1873
Oy 1561 ACCGACAGGCTTGGAGCGGCCACTTACAGGTGGGGAGATGAGACAGATGTCTTCTTA 1620
Db 1874 ACCGACAGGCTTGGAGCGGCCACTTACAGGTGGGGAGATGATACGACGTCTTCTGTC 1933
Oy 1621 TTGAACAGCACTGACACACCGCTGGGTCATGTTGGCTGACCGTGTGATGAACTTCT 1680
Db 1934 CTTAACAAATACAGGACACCGCTGGGCAATTTGTTGGTTTACCTGATGAACTCAACT 1993
Oy 1681 GGGTACACCAAGACTTGGCGGCGCACACCCCTGCGTACTAGAGCTGACTTCAACGCGACG 1740
Db 1994 GGAATTCACCAAGATGTGGAGCGCTCTTGTGTATCGAGAGGCGGGCAACAA-- 2051
Oy 1741 ACGGACTGTTGTGCCCAAGGACTGTTTGAAGAGCATCTGATTAACAATTACCTCAAA 1800
Db 2052 ----CCGTGACGCCCCACTGATTTCTCCGACATCCGAGCCGACATCTCTCGG 2107
Oy 1801 TGGGCTCTGGGCGCTGCTGACAGGCAAGTGTCTGATGACATACCCCTACAGGCTCTGG 1860
Db 2108 TGGGCTCTGGGCTCTGATGACACCAAGTGTCTGATGACATACCCCTACAGGCTCTGG 2167
Oy 1861 CATTAACCTGACAGTTAATTAACATCTTCAAAATAGATGATGTGGAGGGGTT 1920
Db 2168 CATTAACCTGATACATCACTACATTAATTAATCAGATGATACGTGGAGGGGTC 2227
Oy 1921 GAGCAGAGGCTCAAGGCTGATGCAATTTCACTGTGGGAGTCTTGGCACTTGGAGAGC 1980
Db 2228 GAACACAGGCTGGAAGCTGCTCAACTGGACCGGGGGCAAGGTTGCGATCTGGAGAGC 2287
Oy 1981 AGAGACAGAGTCAACTGTCTCTTGTGTGCACTCCACCAAGGAAATGGGCAATTTAACCT 2040
Db 2288 AGGACAGGCTCCAGGCTCAGCCCGTATCTGTACCACTACACAGTGGAGAGTCTCCCG 2147
Oy 2041 TGCCTTACTCGGAGCTGCGCGCTTGTGCACTGTCTTCTCCACCTCCACCAAAATC 2100
Db 2248 TGTTCCTTCAACACCTTACAGGCTTGTCAACGGGCTCATCACTTCCACAGAAATTT 2407
Oy 2101 GTGACGTACAAATTCATGATAGGCTATACCTGACCTCAAAATATACATGTCGATGG 2160
Db 2408 GTGACGTGACATGATGATGAGGGTGGGGTCAAGCATCGGCTCTGGGCAATTAAGTG 2467
Oy 2161 GAGTGGTATTAATCTTATTTCTGCTCTTAAGCGGAGCGCAGGTTTGGCGCTGATAGG 2220
Db 2468 GAGTACGTCGTTCTCTGTTCTTGTGCTTGAAGAGCGCGCTGTGCTCTCTGTTGG 2527
Oy 2221 ATGCTCATCTTGTGGGCGAGCGAGAGCACTAGAGAGCTGTGATCTTGGACGCT 2280
Db 2528 ATGATGCTACTCATATCCCAAGCGAGGCGGCTTTGGAAGAACTCTGTAATCTTAATGCA 2587

2281 GCGAGCGCAGTACGTGCAATGCGCTTCCATATTTTGTGATCTTTTCTGCGCTGCG 2340
2588 GCATCCCGCGCGGAGCGACGCGTCTTGATCTTCCCTGCTTCTTCTGCTTGTGACATGG 2647
2341 TACATCAAGGTGCGGGATGTCCTTATGCTACCTATTCCTCACTGCGCTGTGCTTT 2400
2648 TATTTGAAGGGTAAATGCGGTGCGCGGAGCGGTCTACACCTTCTACGGGAAATGTGCTCTC 2707
2401 AGCTACTGCTCCAGCATTTGCCCCCAAGGCTTATGCGATATGACGATCTGTGCAATGGC 2460
2708 CTCTGCTCTCTTTGGGCTTGCCCCAGCGGCGTACGCGCTGGAACAGAGTGGCCGCG 2767
2461 CAGATAGAGCGCGCTGTCTGTATGATATCACTCTTTTATCTCTACCCCCGGGTATAG 2520
2768 TCGTGTGGCGGTGTGTCTGTGTGGGTGATGCGCGTGACTCTGTACCATATTAACAAG 2827
2521 ACCCTTCTACCGCGTTTTTGTGTGTGTGTGTCTATCTTCTGACCTTGGGGGAAGCTATG 2580
2828 CGCTATATCAAGCTGTGCTTGTGTGTGCTTCAATATTTCTGACCAAGTGAAGCCCAA 2887
2581 GTCCAGAGTGGGACCAACCATATGACAGTGGCGGCGGTGATGTCATATATGGGCC 2640
2888 CTGCACTGTGATATCCCCCTCAACGTCCAGGGGCGCGCGCGCTATCTTATCTC 2947
2641 GTCCCATATTTCAACCAAGGTGTGTGTGTGATACATAACCAAGTGGCTTTGGCGGTCTT 2700
2948 ATGTGTGTGTACCCGCACTGTGTATTTGATATCAACCAATTTGCTGTGCGGTCTTC 3007
2701 GGGCTGTCTTACCTCTTAAAGGTCTTTTGAACGCGCGTGCCTTATCTTGTCAAGGCTCAC 2760
3008 GGAACCCCTTTTGAATTTCTTCAAGCCAGTTTGTCTTAAATGATCCCTTATTTGTGCGGTCAA 3067
2761 GCTCTACTGAGATGTGSCACATATGSCAAGCATATCCCGGGGGGAGATATGTCATATG 2820
3068 GGGCTTCTCGGTTCTGCGGTGATGCGGGAAGTATCGAAGCCATTTACGTGCAATG 3127
2821 GCGCTACTAGCCCTTGGCAGATGATCGGCATTTACATATGACATCAACCCCTATG 2880
3128 GTCATCTTAAATTTAGGGGCGCTTATCTGSCATCTATTTTATACATCTCACCTCTT 3187
2881 TCGATTTGGCTGTAGTGCCTGCGGAGCTTGGCGGTGCGCTTGAAGCTTATCATTTTC 2940
3188 CGGAGCTGGGCGCAACACGGCTTGGCGAGATCTGGCGTGTGATGAGCCAGTGTCTTC 3247
2941 AGTCCCATATGAGAAAGATCATTTGTGTGGGAGCGGAGACAGTGTGTGGGGAAT 3000
3248 TCCCAATATGAGACCAAGCTTATCACTGTGGGAGCATACCGCGGTGCGGTGATC 3307
3001 TTAACAGGACTTCCCGTGTCCGCCGACTTGTGTGGAGGTCTCTTGGCCCACTGAT 3060
3308 ATCAACGGCTTGTGCTGTTCGCCCGGAGCGGAGGCGGAGATCTGTCTGGGCGCACCGAT 3367
2881 TCGATTTGGCTGTAGTGCCTGCGGAGCTTGGCGGTGCGCTTGAAGCTTATCATTTTC 2940
3188 CGGAGCTGGGCGCAACACGGCTTGGCGAGATCTGGCGTGTGATGAGCCAGTGTCTTC 3247
2941 AGTCCCATATGAGAAAGATCATTTGTGTGGGAGCGGAGACAGTGTGTGGGGAAT 3000
3248 TCCCAATATGAGACCAAGCTTATCACTGTGGGAGCATACCGCGGTGCGGTGATC 3307
3001 TTAACAGGACTTCCCGTGTCCGCCGACTTGTGTGGAGGTCTCTTGGCCCACTGAT 3060
3308 ATCAACGGCTTGTGCTGTTCGCCCGGAGCGGAGGCGGAGATCTGTCTGGGCGCACCGAT 3367
3061 GGGCTATATCTTCAAGGGGTGAGTCTTCTGCGCCCATCACTGTCTTACCGCCAGACAGA 3120
3368 GGAATGGTCTCAAGGGGTGAGAGTGTCTGCGCCCATCACTGCGGTGATGAGCCAGACAGA 3427
3121 CGTGGCTTTTGGGACCATATGTGTGATGATGACGCGGCGGACAGAAAGAGGCT 3180
3428 AGGGGCTCTCTAGGGTGCATATCAACAGCTTAACTGTGCGGAGCAAAAACCAAGTGGAG 3487
3181 GGGGAAATTTAGGCTCTGTCCACAGTCACTCAAGCTTCTCTCGGACATCATCTGGGG 3240
3488 GGTGAGGTCCAGATGTGTGATCACTGTGCGCCAAACCTTCTTGGGACGTGTGATATGGG 3547
3241 GTTTTGTGACTGTCTTACATATGAGCTGTGCAACAAAGCTTGTGCGGCTACAGGAGTCCG 3300
3548 GTGTGTGAGACTGTCTTACAGGGGCGGAAACGAGACCAATGCGGTACACCAAGGTCCT 3607
3301 GTCAAGAGATGATCTCAAGTGTGAGGGGAGCTTATATAGGTGGCCCAAGCCCTCCGCG 3360
3608 GTATCTCAAGATGTATCAATGTATGACCAAGACCTTGTGTGGCTGTGCGCCGTCCGCAAGT 3667
3361 ACTAAATCTTTGAGCGGTGACAGTGTGAGCGGTGACCTGTATCTGTATCAGCGGAC 3420

3668 AGCGGCTCATTTGACACCTCGACATGTGGCGCTCCCGGACCTTTATCTGTATCAGAGGCAC 3727
3421 GCTGATGTATTCGGGCTCGAAGACGCGGAGCAACCGGAGGCTATCTTCCCGAGA 3480
3728 GCGGATGTATTCGGGCTCGGCGGCGGTGATAGCAGGGGAGCGCTGTGCTGCGCCGG 3787
3481 CCTTTTGCACCTTGAAGGGGTCTCTGAGAGGCGCGGTGCTATGCGCCAGGGGCGAGCT 3540
3788 CCAATTTCTTATTTGAAGGCTCTCTGCGGAGGTCTGCTGTGTGTCGCCCGGCGGACGCG 3847
3541 GTCGAGTCTTCCGGGAGCTGTGTGTCTTCCGGGCGTGTGCTATGATCATATGATTCATC 3600
3848 GTGGGATATTTAGGGCGCGGTGTGTGACCGGTGATGAGTGTGCTAAGGGGTGATATC 3907
3601 CCGGTGAGACCTCGACATGTGACGCGGTCCCGGACCTTTTATGACAAACAGACACA 3660
3908 CTTGTGAGAACTTGAAGACCAACATGAGGTCCCGGTGTTCACGATTAATCTCTTCCA 3967
3661 CTTGTGTGCGCCGAGACTTATCAAGTGTGAGTCTTGTGATCCCGGACCTGGCAGTGAAG 3720
3968 CAGTATGTCCCGAGAGCTTCAAGTGTGCTCACTCTCATGTCTCCACAGGAGCGGAGAA 4027
3721 AGCACAAGTCTCTGTGATATGTGTCTCAAGGGTATAAAGTGTATGCTTATATCC 3780
4028 AGCACAAGTCTCTGTGATATGTGTCTCAAGGGTATAAAGTGTATGCTTATATCC 4087
3781 TCAGTGTGCTCACCTGTGGGCTTTGGGCGTATCTTGTATGAGCACATGCAATATCC 3840
4088 TCTGTGTGCTCAACACTGTGGCTTTGTGTGTATATGATCCAAAGCTCATGTGATGATCT 4147
3841 AACATTAAGATCTGAGTCAAGATCTGAGTCAAGCGGCGGCGGCACTCATGCTCCACATAT 3900
4148 AACATTAAGATCTGAGTCAAGATCTGAGTCAAGCGGCGGCGGCACTCATGCTCCACATAT 4207
3901 GGCATAATCTCTCGGCTTGGGCGGTGTGCGGCGGCGGCTTACAGACATCATATATGAT 3960
4208 GGCAGTGTCTTGTGCGGCGGCGGTGTGCGGCGGCGGCTTATGACATTAATTTGTGAC 4267
3961 GAAATGCCATGCGGTGATCTTACCAACATCTTGGATGCGGAGAACATGTCTTGTATCAAGCA 4020
4268 GAGTGTGCTCAACAGATGACATCATCTTGGGATCTGCGCTGTGACCAAGCA 4327
4021 GAGACAGTGTGGGTGACATCACTGTGTGCTGCTTACAGCTACGCGGCTGTGATGATCA 4080
4328 GAGATGTGCGGCGGAGT 4387
4081 ACCCGGCGGCGGAGATGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4140
4388 GTGCGGCTTCCCAACATGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4447
4141 GGGAGGCGGATTTCCCTGTCTTATCATATGAGGAGAGAAAGATCTTGTGTGATCAATCA 4200
4448 GGCAGGCTTATCCCTCGAAGTATCAAGGGGAGAGATCATCTTCTGTGATCAATCA 4507
4201 AAGAAAAGTGTGAGCAAGCTGTGCGGCGGCGGCTTGTGCGGTATGAGCTTGAATCAATG 4260
4508 AAGAAAAGTGTGAGCAAGCTGTGCGGCGGCGGCTTGTGCGGTATGAGCTTGAATCAATG 4567
4261 TACTAGAGAGGTTGTGAGCTGTCCGTAATACCACTCAAGGAGAGAGTGTGTGTGTGTGTGT 4320
4568 TACTAGAGAGGTTGTGAGCTGTCCGTAATACCACTCAAGGAGAGAGTGTGTGTGTGTGTGT 4627
4321 ACCGAGCGGCTCATGACAGGTTATCTGAGGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 4380
4628 ACCGAGCGGCTCATGACAGGTTATCTGAGGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 4687
4381 GGGGTACATCAAGTGTGAGCTTCAAGTGTGAGCTTCAAGGAGAGTGTGTGTGTGTGTGTGTGT 4440
4688 TGTGTACCAAGAGTGTGATTTTCAAGCTTGTGAGCTTCAAGGAGAGTGTGTGTGTGTGTGTGT 4747
4441 GTCCCTCAAGAGCTGT 4500

D 4748 CTCCTCCAGGATGCTGTCTCCCGCATCAACGTCGGGAGAGACTGACAGGGAGAACCA 4807
Q 4501 GGCAATTAAGATGATGTTTCCATGTAAGAGAGCTCAGGAATGTTGACAGTATAGT 4560
D 4808 GGCAATTAAGATGATGTTTCCATGTAAGAGAGCTCAGGAATGTTGACAGTATAGT 4867
Q 4561 CTCTGTGATGCTACGACGAGGAGGCGCATGTATGAGCTCACACCATGAGAGACACC 4620
D 4868 CTCTGTGATGCTACGACGAGGCTGTGTATGAGCTCACACCATGAGAGACACC 4927
Q 4621 GTCAGGCTCAGGGGATATTTCAACACCGCGGTTGCTGTGTGCGAAAGCATTTGAG 4680
D 4928 GTTAGGCTACGAGGCTACATGAACACCGCGGCTTCCGTGTGCGAAAGCATTTGAG 4987
Q 4681 TTTTGGAGGAGATTTTCAACACCGCGGCTCAGACATAGTCCCATTTCCCAACA 4740
D 4988 TTTTGGAGGAGGCTTCAACACCGCGGCTCAGACATAGTCCCATTTCCCAACA 5047
Q 4741 AAGCAATCGGGGAAAAATTTGCACTTAAACAGCTTACAGGCTTACAGTGTGCTAGG 4800
D 5048 AAGCAATCGGGGAAAAATTTGCACTTAAACAGCTTACAGGCTTACAGTGTGCTAGG 5107
Q 4801 GCCAAAGCCCCCCCCCGCTGCTGGAGGCTCATGTGAAAGTTTGAATTCGATCAACACCC 4860
D 5108 GCTCAGGCCCCCCCCCGCTGCTGGAGGCTCATGTGAAAGTTTGAATTCGATCAACACCC 5167
Q 4861 ACACCTGTGGGCCCCCAGCATCTCTCTGTACCGCTTGGGCTCTGTAAACAGAGTACCC 4920
D 5168 ACCCTTCATGGGCGCAACCCCTGTCTATACAGACTGGGCGCTGTTCAGAAATGAATACCC 5227
Q 4921 CTCACACATCCCGTGAAGAAATACATGAGCACTGACATGCAAGCCGACCTTGAAGTATG 4980
D 5228 CTGACGACCCAGTACCAAAATACATGATGATGAGGAGAGTCTCTCAGCACTTACCG 5287
Q 4981 ACCGACATGGGCTTGGGAGGAGGCTTGGGCGGCTGCGCGGCTTATGCTGGGCG 5040
D 5288 ACCGACATGGGCTTGGGAGGAGGCTTGGGCGGCTGCGCGGCTTATGCTGGGCG 5347
Q 5041 ACCGAGTGTGTTTTCATCATCAGCGCTTGGGCTTGGGCTTATGCTGGGCG 5100
D 5348 ACAGGCTGCGTGTCTATGATGAGGAGGCTCTTGTGCGGAGAGCGGCAATCATACCT 5407
Q 5101 GACAAAGAGTCTCTATGAGGCTTGTGATGATGAGAGATGCTCTTGAAGGCGCT 5160
D 5408 GACAGGAGAGTCTCTACCGAGAGTGTGATGATGAGAGAGTCTCTCAGCACTTACCG 5467
Q 5161 CTCATTGAAGAGGAGGAGGCTATGAGGAGATGCTGAAGTCCAGATCCAGAGCTTATG 5220
D 5468 TACATCGAGCAAGGAGATGATCTGCGGAGAGTTCAGAGAGAGGCTCTGAGCTCTG 5527
Q 5221 CAGCAAGCTTCCAAACAGCTCAAGACATACCAACCACTGTGACAGGCTTACAGCCAG 5280
D 5528 CAGACCGGCTCCGTCAGGAGAGGTTATGCGCCCTGTGTCTCAGACCAATGCGAAGAAA 5587
Q 5281 GTAGAACATTTCTGGGCAACACATGTGAACTTCAATAGCGGATCCAAATACCTGCA 5340
D 5588 CTCAAGACCTTCTGGGCAACACATGTGAACTTCAATAGCGGATCCAAATACCTGCA 5647
Q 5341 GGACATCAACATGCGCAGGAGACCTGAGATGCTTCATGATGAGGCTTTCACAGCTGCT 5400
D 5648 GGCTTGTACAGCTGCTGTGTATCCCGCATTTGCTATGATGAGGCTTTCACAGCTGCT 5707
Q 5401 CTCACAGTCCGCTGTCAACAGACACATATCTTCAACATTTTGGGGGCTGTGCTA 5460
D 5708 GTCAACAGCCCACTAACCACTAAGCAAAACCTCTCTTCAACATATTTGGGGGCTGTG 5767
Q 5461 GCATCCCAATTTGCAACACCGCGGGGCACTGGCTTCTGTGTAGTGTGCTTATGAGGA 5520
D 5768 GCTCCCGGCTGCGCGCGCGGTGCTCTACTGCTTGTGTGCGGCTGTGAGTGTGCG 5827
Q 5521 GCTCCCGGAGCATATGAGGCTTATGATGATGAGTGTGAGTGTGAGTGTGAGTGTGAGT 5580
D 5828 GCTCCCGGAGCATATGAGGCTTCTCATAGACATCTCTTGTGCGGCTGTATGCG 5887

Q 5581 GCGGAGATTTGGGGGCTCTCGTGCATTCAGATCATGTGTGCGAGAGAGCTTCATG 5640
D 5888 GCGGAGGCTGGGGAGCTCTTGTGCAATTCAGATCATGAGGAGGAGGCTCCCTCCAG 5947
Q 5641 GAGGATGTGTCACTTGTGCTGCTGGAATTTCTGTCTCGGGGCTTGTGTAGTGGAGTC 5700
D 5948 GAGGACCTGTCAATCTACCTGCGGCAATCTCTGCGCGGAGGCTCTCTATGCGGCTG 6007
Q 5701 ATCTGCGGCGCATTTGCGCGGCAACGTTGGGAGCGGGGAGAGGCGCTCCCATAGATG 5760
D 6008 GTCTGTGACCAATTTCTGCGCGGCAAGTTGGCGCGGAGGAGGAGGAGTGTGAGT 6067
Q 5761 AATGACATGATGCTTTTCTTCCAGAGGAAATCAAGTGTGCGGCGGCACTTACGTTAG 5820
D 6068 AACCGGCTATGCTTGTGCTCTCCGCGGAGAACATGTTTCCCGCAGCATACGTCGCG 6127
Q 5821 GAGTCCGATGCTGTGAGGCTGTGACCCCACTACTTGTGCTCCCTTACATACCAAGCTG 5880
D 6128 GAGGAGATGCAAGCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6187
Q 5881 CTCAGAAAGCTCCCACTGATTAATGAGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 5940
D 6188 CTGAGGAGCTGCAACGATGATGATGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGG 6247
Q 5941 CTCGCGATGTGTGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6000
D 6248 CTGAGGAGATCTGAGGATGATGATGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGG 6307
Q 6001 TCCAAATTAATCCCAAGATGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6060
D 6308 GCTTAACTATGCAAGCTGCTGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6367
Q 6061 GCGGCTGTGCGGCACTGAGCATCATGACCAACAGGATGCTTGTGCGGCGGCAATATCT 6120
D 6368 GCGGCTGTGCGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAG 6427
Q 6428 GGCATGTCAAAAACGAGAGATGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGT 6487
D 6181 CAGGAGCTTCTCTATCAATTTGTTACAGGAGGCTGAGTGTGAGGAGTGTGAGGAG 6240
Q 6488 AGTGGAGCTTCCCAATTAATGCTTACCAACAGGAGGCTGAGTGTGAGGAGTGTGAG 6547
D 6241 AACTTAAGTGTGCACTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAG 6300
D 6548 AACTTAAGTGTGCACTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAG 6607
Q 6301 GAGTCAATCCATCACTAACAGGACTCACCACTGATTAATTAATGCTTCCCTGCACTA 6360
D 6608 GGGAGCTTCCCACTAACAGGACTCACCACTGATTAATTAATGCTTCCCTGCACTA 6667
Q 6361 CCTCTCCGAGTCTTCTTCTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTG 6420
D 6668 CCATGCGCGGAATTTTCAAGAAATTTGAGGAGGAGTGTGAGGAGTGTGAGGAGTGTG 6727
Q 6421 CCGAAGCGTCTTCTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 6480
D 6728 TGAAGGCTTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTG 6787
Q 6481 GAGTCCAGGCTTCTTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGG 6540
D 6788 GAGTCCAGGCTTCTTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGG 6847
Q 6541 GATTCATTCATTAACAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 6600
D 6848 GATTCATTCATTAACAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 6907
Q 6601 TCCGAGGAGGCTTCTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 6660
D 6908 TCTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 6967

OY	6661	A	CCGACGGCAAAAGCCTAATGATGTGACACTGGTGAATGCTAACTT-----GTTCC	6708
Db	6968	G	TAAACATGACTCCCTGATGTGAGGTCTATAGAGCCAACCTCTATATGAGGCGACAGAG	7027
OY	6709	A	TGGGGGGGCGATGTGACTCCGGAATAGAGCTGGAGTCCAAAGAGTGTCGTTCTGAGACTCTTC	6768
Db	7028	A	TGGGGGGGCAACATCAACGAGGTGATGTCAGAAAAAAGATGGATTTCTGAGCTCTTTC	7087
OY	6769	G	ACCCCAATGTGTCGAAGAAAGAGCGACCTTGAGCCTTGATACCATCAAGATPACATGCTC	6828
Db	7088	G	ATCCGCTTGTGGCGGAGAGGAGCGAGGAGATCTCCGTAACCCGACGAATCTTGCGG	7147
OY	6829	C	CCCAAGAAAGAGTGTCCCAACAGCTTTACCGGCTGTGGGACAGGCTGATTAACAACCCACG	6888
Db	7148	A	AGTCTCGGAATTCGCCAGGCCCCGTGCGCTTTGGGGGCGGCGGACCTATAACCCCCCG	7207
OY	6889	C	TGTGGAATGATGGAAAAAGCGAGATTACCAACCGGCGACCTGTGTGGGGGTGGTCTC	6948
Db	7208	C	TATGTTGAGAGCTGGAAAAAGCCGACCTACCAACCACTGTGTGTCATGCTGTGCTT	7267
OY	6949	C	CTCTCTCTGAGAAAAACCCCGACGCTTCCCAAGAGAGCGCGACAGTGGGCTAAGT	7008
Db	7268	C	CAACTCCAAAGTCCCTCTGTGTGCTCTCGCTCGGAAGAGCGAGGTGTCTCACT	7327
OY	7009	G	AGGACTTCATAGAGAGATGCCCTTCAACAGCTGGCTAATTAAGTCTTTGGCAGACCCCC	7068
Db	7328	G	AATCAACCCCTATCTACTGCTTGGCCAGGCTCGCACACAAAGCTTTGGCAGCTCTCA	7387
OY	7069	C	CAAGCGGCGAATTCAGGCTTTCCACGGGGGGCGGGCGGTGCCAATTCGGGAGTCAAGAG	7128
Db	7388	A	CTTCCGACATTAACGGGCGACAAATACGAACAATCTCTTGAGCCGCGC-----	7435
OY	7129	C	CTCTCTGATGAGTTGGCCCTTTCCGAGACAGGTTTCATCTTTCATAGCCCCCTCGAG	7188
Db	7436	C	CTTCTGGCTGCTCCCCCGGACTCGGACGCTGATGTCTATTCTCTCAATGCCCTCGTAG	7495
OY	7189	G	GGGAGCTTGGAGATCCAGACTGGAGCTGAGCGGTAGAGCCCCCAACCCCCCCACG	7248
Db	7496	G	GGGAGCTTGGAGATCCGGA-----	7515
OY	7249	G	GGGGGGGTGGAGCTCCCCGCTCGGACTCGGGGTCTGTGTCTACTGTCTCGGAGAGAC	7308
Db	7516	-----	CTTTAGGACGGGGCATGTGTCAACGATCAATGATGAGGCGCAACGCG	7561
OY	7309	G	ACTCCGTCGATGTGCTGCTCCATGTCATATCTCTGACCGGGGGCTTAATTAATCTCTGT	7368
Db	7562	G	AGGATGTCTGTGTGCTCTCAATGTCTTACTCTTTGGAGAGCGGACTCTGTCAACCCGTGC	7621
OY	7369	A	GTCCCGAAGAGAGAAAGTTAACCGATTAAACCTCTTGAGCACTCCCTGTTGGCATATCAC	7428
Db	7622	G	CGCGGGAAGAAACGAAGATCGCCATCAATGCATTAAGCAACTGTGTGTACATGTCACAC	7681
OY	7429	A	ACCAAGTGTACTGTACCAACAACAAGAGCGCTCATCTAAGGGCTAAAAAGTAACTTTT	7488
Db	7682	A	ATTGTGGTATTCACCACTCAACGACTCTTGCCAAAGCGAAGAAAGTCAACTTT	7741
OY	7489	G	ATGATGATGACAGTGTGACCTGCTCACTACAGACTCAGCTTAAAGGACTAATGACTAGCG	7548
Db	7742	G	ACGACACTGCAAGTTCTTGACAGACCATTAACAGACGACTCAAGAGAGTTAAAGCAGCG	7801
OY	7549	G	CGCTTCAAGATCAACCGCAAGGCTCCTCAACATGAGAGAGGCTTGCCAGTTAAACCCACCC	7608
Db	7802	G	CGTCAAAAGTGAAGGCTTAACCTTGATTCGTAGAGGAAGTTGCAGCTGAGCGCCCCA	7861
OY	7609	C	ATTCTGCAAGATCTAAATATAGGTTTGGGCTTAAGAGGTTCCGACGTTGTCCGGAGAG	7668
Db	7862	C	ACTCAAGCAATCCAAAGTTGTGTTATGGGCAAAAGCGTCCGTTGTCATGCTCAGAAAG	7921
OY	7669	G	CGCGTTAACCAATCAAGTCCGTGTGGAAGAGCCCTGAGAGAGCTAGCAAGAACCAATT	7728
Db	7922	G	CGCGTTAACCAATCAATCTCGTGTGGAAGAGCTTTCTTGGAGACAAATGTAACCAATA	7981
OY	7729	C	CCCAACCATTAATGCGCAAAAAAGAGGTGTTCTGCTGACCCCAACCAAGGGGGGCAAG	7788

Db	7992	GACACTACCAATCAGGCTTAAGAAAGAGGATTTCCTGCGCTTACGCTTCGAGAAAGGGGCTCT	80411
Qy	7789	AAAGCAGCTCCCTTATGTTTACCTGACCTCGGCGTCAGGGTCTGGAGAAATGGCC	78448
Db	8042	AAGCAGCTGCTGCTCATGCTGTTCCTCCGATCTGGGCGCGCGCTGTGGGAAAGATGGCT	81010
Qy	7849	CTTATGACATTAACA AAAA CTTCTCTCAGGCGGTGATGGGCGCTTATGATTCAG	79080
Db	8102	TTGTATCGACGTGGTTACAAAGCTCCCTTCCTGGCCCTGATGGGAAAGCTCTTACGGATTCCAA	81610
Qy	7909	TATTCCCCCGCTCAGCGGCTGAGATTTCTCTTGAAGATGGGCGGAAAGAAAGAACCTT	79668
Db	8162	TACTACCAAGACAGCGGCTTGAATTCCTCTGCAAGGTGGAAGTCCAAAGAAACCCCA	82211
Qy	7969	ATGGGTTTTTGTATGATATACCCGATNGCTTTTACTCAACCTGCTACGTGAGAGACATCAG	80238
Db	8222	ATGGGTTCTGTATGATATACCCCTGCTGTTTACTTCACAGTCACTGAGACGACATCCGT	82818
Qy	8029	ACTGAGAGTCCATATATCGGGCCCTGCTCTTTCGCGAGAGGCCCACTGCCATATAC	80888
Db	8282	ACGGAGAGGCAATCTACCAATGTTGTGACCTTCGACCCCCAAAGCCCGGTGGCCATCAG	83411
Qy	8089	TCGCTAATGAGACATTTACGTGGAGGGCTTATGTTCAAACAGCAGGGCCAAACTGTC	81488
Db	8342	TCCTTCACCGAGAGGCTTATGTTGGGGGCGCTTTACCAATTCAGGGGGGAAACATGCG	84010
Qy	8149	GGGTACAGGCGTGGCGCGCGCAGGCGGGGTGCTCACTACCTAGCATGGGGAAACACATCACA	82080
Db	8402	GGCTATCGCAGGTCCGCGCGAGCGGCGTACTGACAACTAGCTGTGGTAACCTCTACT	84610
Qy	8209	TGCTACGTGAAGCCTTAGCGGCTTGTAAAGCTGCAGGGATATATCGCGCCCAATGCTG	82688
Db	8462	TGCTATCATCAAGGCCCGGGCGAGCCTGTCTGACCGCAGAGGCTCCAGGACTGACCATGCTC	85210
Qy	8269	GTAATCGGCGGATGACTGTGTGTATCTCAGAAAGCCAGGGGACCGAGAGAGACGAGCG	83288
Db	8522	GTGTGTGCGACGACTTATGCTTATCTGTGAAAGCCCGGGGGTCTCAGAGAGGACGCGCG	85810
Qy	8329	AACTGAGAGCTTACAGGAGGCTATGACAGAGATTTCTGCCCCCTCGGTGGAACCCCCC	83888
Db	8582	AGCTGAGAGCTTTCAGGAGGCTATGACAGAGTATCTCGCCCCCTCGGGGAAACCCCCA	86411
Qy	8389	AGACCGAGTATGATCTGAGCTGAGCTGATAACATCTTGTCTCTCAAAATGTTGTGTGGCGCTG	84488
Db	8642	CAACCAAGATACGACTTGGAGCTCATTAACATCACTGCTCTCCAAACGTGTCAAGTGGCCAC	87010
Qy	8449	GGCCCAACAGGCGCGCCGACATTAATCTTACCTGACAGAGACCTTACCACTCCATTCGCCCG	85080
Db	8702	GACGGCGCTGAAAGAGGGGTCTACTTACTTACCTCACCGTGAACCTTACAAACCCCTCGCAGGA	87610
Qy	8509	GCTGCTTGGGAAACAGTTAAGACATCTCCCTGTCAATTCATGAGCTGGGAAACATATATCAG	85668
Db	8762	GCTCGTGGGAGACAGCAAGACACATCTCAGTCAATTCCTGGCTTGAAGCAATATATCATG	88210
Qy	8569	TACGCCCCGACATATGAGGCTCGCATGCTCTGTATGACACACTTCTTCTCATCTCATG	86288
Db	8822	TTTGCCCCCACTGTGTGGCGAGGATGATATGATGATCCATTTCTTTAGGTCCTTATA	88810
Qy	8629	GCTCAAGACAGCTGAGACAGAACTTCACTTGGATGTACGAGCGGCTGACTCCGCTG	86888
Db	8882	GCGAGGAGCCAGCTTGAACAGGCGCTCGATTGCGAGATCTACGGGGCGCTGCTACTCATATA	89411
Qy	8689	AGTCCCTTGAACCTCCCACTATATTAATGAAAGTTAATAGGCTTGAACGCTTTTCTCTG	87488
Db	8942	GAACCACTTGTATCTTCACTCATATATTCAAAGACTCAGTGCCTTCAGCGCATTTTCACTC	90010
Qy	8749	CACACATATCACTCCCAAGAACTGACACAGGGTGGCTTCAGGCCCTCAGAAAACTTGGGCG	88080
Db	9002	CACAGTATCTTCAAGGTGAATTAATATAGGTGGCCGCAATGCTCAGAAAACTTGGGGTA	90610
Qy	8809	CCACCCCTCAGAGCTGGAAGAGCCGGGCACTGTCAGTCAAGGCGTCTTCATTCCTCGT	88668

Db 9062 CCGCCCTTGCGAGCTTGAGACACCGGCCCCGAGCGCTCCGCGTAGCGCTTGTGGCCAGA 9121
QY 8869 GGGGGGAGAGCGCGCGCTTTGGGTGATATCTTTCATTGGGCGGTGAAGACCAAGCTC 8928
Db 9122 GGAGGGAGGGCTGCATATATGGAAGTACTTTCACACTGGGAGTAAGACAAAGCTC 9181
QY 8929 AAATCACTCATTTGCCGGAAGCGCGCTCTGTGATTTATCCAGCTGTTCACCGTCGCG 8988
Db 9182 AAATCACTCAATAGAGCGCGCTGCGCAGCTGTGACTGTTCGGCTGTTCACGCGTGGC 9241
QY 8989 GCGGCGGGGCGACATTTATACAGCGGTGTCGCGTCCCGACCGCTATTGCTCTTT 9048
Db 9242 TAGAGCGGGGAGACATTTATACAGCGGTGCTCATGCGCGCGCTGATCTGTTT 9301
QY 9049 GCGCTACTCTACTTTTGTAGAGGGTAGGCGCTTTCTACTCCCGCTCG 9098
Db 9302 TGCTTACTCTGCTGTGCGAGGGGTAGGATCTACTCTCCCAACCG 9351

RESULT 15
US-07-910-760-9
: Sequence 9, Application US/07910760
: Patent No. 5683864
: GENERAL INFORMATION:
: APPLICANT: Houghton, Michael
: APPLICANT: Choo, Qui-Lim
: TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
: TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Chiron Corporation
: STREET: P.O. Box 8097 (Int. Prop. R-440)
: CITY: Emeryville
: STATE: CA
: COUNTRY: U.S.A.
: ZIP: 94662-8097
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/910.760
: FILING DATE: 07-JUL-1992
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Blackburn Bsq., Robert P.
: REGISTRATION NUMBER: 30,447
: REFERENCE/DOCKET NUMBER: 0101.002
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (510) 601-2702
: TELEFAX: (510) 655-3542
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 9401 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 342..9374
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 366
: OTHER INFORMATION: /note= "This amino acid position
: OTHER INFORMATION: can also be Arg."
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 372
: OTHER INFORMATION: /note= "This amino acid position
: OTHER INFORMATION: can also be Thr."

: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 867
: OTHER INFORMATION: /note= "This amino acid position
: OTHER INFORMATION: can also be Thr."
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1341
: OTHER INFORMATION: /note= "This amino acid position
: OTHER INFORMATION: can also be Val."
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 2148
: OTHER INFORMATION: /note= "This amino acid position
: OTHER INFORMATION: can also be Ile."
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 2883
: OTHER INFORMATION: /note= "This amino acid position
: OTHER INFORMATION: can also be Asn."
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 3681
: OTHER INFORMATION: /note= "This amino acid position
: OTHER INFORMATION: can also be Ser."
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 3690
: OTHER INFORMATION: /note= "This amino acid position
: OTHER INFORMATION: can also be Thr."
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 4167
: OTHER INFORMATION: /note= "This amino acid position
: OTHER INFORMATION: can also be Leu."
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 4323
: OTHER INFORMATION: /note= "This amino acid position
: OTHER INFORMATION: can also be Val."
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 4701
: OTHER INFORMATION: /note= "This amino acid position
: OTHER INFORMATION: can also be Tyr."
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 4752
: OTHER INFORMATION: /note= "This amino acid position
: OTHER INFORMATION: can also be Ser."
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 5970
: OTHER INFORMATION: /note= "This amino acid position
: OTHER INFORMATION: can also be Gly."
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 6183
: OTHER INFORMATION: /note= "This amino acid position
: OTHER INFORMATION: can also be His."
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 6186
: OTHER INFORMATION: /note= "This amino acid position
: OTHER INFORMATION: can also be Cys."
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 6402
: OTHER INFORMATION: /note= "This amino acid position
: OTHER INFORMATION: can also be Val."
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 7386

OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7494
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7497
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ala."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7845
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8409
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9102
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9327
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Pro."
OTHER INFORMATION: can also be Pro."
US-07-910-760-9

Query Match 45.2%; Score 4115.2; DB 1; Length 9401;
Best Local Similarity 66.3%; Pred. No. 0;
Matches 6042; Conservative 0; Mismatches 2978; Indels 90; Gaps 5;

QY 1 ATAGCAAAATCTTAACCTCAAGAAACCAAAACCAACCAACCGTCCGCCA 60
DB 342 ATAGCAAGATCTTAACCTCAAGAAACCAAAACCAACCAACCGTCCGCCA 401
QY 61 GAGCTTAAGTTCCGGGCGGCGGCGAGATGTTGGCGAGATATCTTTGGCGGAGG 120
DB 402 GAGCTCAAGTTCCGGGCGGCGGCGAGATGTTGGCGAGATATCTTTGGCGGAGG 461
QY 121 GGGCCGAGTTGGGTGTGCGCGGCAAGAGAACTTCGAGCGGTCACCGACGTGA 180
DB 462 GGGCCGAGTTGGGTGTGCGCGGCAAGAGAACTTCGAGCGGTCACCGACGTGA 521
QY 181 AGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
DB 522 AGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 581
QY 241 TACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
DB 582 TACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 641
QY 301 CGAGGTTCCGCTTCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
DB 642 CGAGGTTCCGCTTCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 701
QY 361 AAGGTATGATACCTTAAGTGTGGGTTGGCGGCGGCGGCGGCGGCGGCGG 420
DB 702 AAGGTATGATACCTTAAGTGTGGGTTGGCGGCGGCGGCGGCGGCGGCGG 761
QY 421 GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
DB 762 GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 821
QY 481 GGGGTAAATTTTGAACAGGAACTTACCGGTTGCTCTTTTCTATCTTTGCTG 540
DB 822 GGGGTAAATTTTGAACAGGAACTTACCGGTTGCTCTTTTCTATCTTTGCTG 881

QY 541 CTGCTGCTGCAATCAACCCCGGCTCTCGGCGGAGTGAAGACATCAATACCGG 600
DB 882 CTGCTGCTGCTGCAATCAACCCCGGCTCTCGGCGGAGTGAAGACATCAATACCGG 941
QY 601 TACATGATGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 660
DB 942 TACATGATGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1001
QY 661 CTGCAAGTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
DB 1002 CTGCAAGTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1061
QY 721 CCGGTCTACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 1062 CCGGTCTACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1121
QY 781 CACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 1122 CACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1181
QY 841 TGGGTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 1182 TGGGTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1241
QY 901 GTCCAGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 1242 GTCCAGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1301
QY 961 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 1302 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1361
QY 1021 CCGAGGTCAATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1080
DB 1362 CCGAGGTCAATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1421
QY 1081 TACTTCTATGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
DB 1422 TACTTCTATGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1481
QY 1141 GTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
DB 1482 GTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1541
QY 1201 AGCTTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1542 AGCTTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1601
QY 1261 CACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1602 CACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1661
QY 1321 CTGTTTCAACCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380
DB 1662 CTGTTTCAACCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1721
QY 1381 AGTATGAGGCTTCCGGGTGGATGATGATGATGATGATGATGATGATGAT 1440
DB 1722 CCGCTTACCATTTTGAACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1779
QY 1441 CAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1780 CAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1855
QY 1501 GCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
DB 1856 GCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1895
QY 1561 ACCGACAGGCTTGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1620
DB 1896 ACCGACAGGCTTGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1955

Qy 1621 TTGAACAGCACTGACACCGCTGGGGATGATGTTTCGGCTGACAGTGGATGAACTCTTCT 1680
Db CTTAAACAATACCAAGCCACCGCTGGGCAATTGGTTGGTATACCTGATGAACTCAACT 2015
Qy 1681 GGCCTACACCAAGACTTGGCGGCGACACACCTGCGCTACTAGAGCTTCAACGCCAGC 1740
Db 2016 GGATTTCACAAAGTGTGGGAGCGCTCTTGTGTCACTCGAAGGGGGGCGCAACAA- 2073
Qy 1741 ACGGAACCTGTTGTGCCCAAGGACTGTGTTTGAAGAGATCCTGATCACTTACCTCAAA 1800
Db 2074 ----CCCTGACACTGCCCACTGATTTGCTTCGCAAGCATCCGACGCACTACTCTCGG 2129
Qy 1801 TGGGCTCTGAGGCTTGGCTTCAAGCAAGTGGCTGATCACTACCCCTTACAGGCTTGG 1860
Db 2130 TGGCGCTCCGGTCCCTGGATCAACACAGGTGGCTGGTCACTACCGGTATAGGCTTTGG 2189
Qy 1861 CATTAACCCCTGCAAGTTAACTATACCATCTTCAAAATPAAGATGATGTGGAGGGGTT 1920
Db 2190 CATTAATCCTGTACTCACTCACTACCAATTTAAATCAAGATGTACGTGGAGGGGTC 2249
Qy 1921 GAGCACAAGGCTCAGCGCTGCACTGCAATTTTCACTGCTGGGATCGTTGCAACTTGGAGAC 1980
Db 2250 GAACACAAGGCTGGAAGCTGCTGCACTGACGCGGGCGAAGCTTGCATCTGGAAGAC 2309
Qy 1981 AGAGACAAGATCAACTGTCTCTCTTTGTGCACTTCAACAGGAATGGGCCATTTTACT 2040
Db 2310 AGGACAAGTCCGAGCTCAGCCCGCTACTGCTGACACTACAGTGGCAGGTCTCTCCG 2369
Qy 2041 TGCCTTAATCTGGACCTGCCCGCTTGTGCACTGCTTCTTCCACTCTTCCACCAAAATATC 2100
Db 2370 TGTTCCTTCAACACCTTCAACGACCTTGTTCACGCGGCTCATCCACTCCACCAAGAACTT 2429
Qy 2101 GTGACGCTACAACTCATGTATGAGCTTATCACTGCTCCCTCACAAATATCATGTCCGATG 2160
Db 2430 GTGACGCTGCACTGTATGTAGGGGTGGGGTCAAGCATCGGCTCTGGGCAATTAAGTGG 2489
Qy 2161 GAGTGGGTAACTCTTATTTCTGCTCTTTCAGGAGCGGCAAGGTTTGGCTGTATGG 2220
Db 2490 GAGTACGTCGTTCTCTGTTCTTCTGCTTGCAGACGCGGCTCTCTCTGCTTGTGG 2549
Qy 2221 ATGCTCATCTTGTGGGCGAGCGGAGCGACACTTAAGAAAGCTGTGCATCTTGCAGAGCT 2280
Db 2550 ATGATGCTACTCATATCCCAAGCGAGCGGCTTGGAGAACTCGTAATACTTAAGCA 2609
Qy 2281 GCGAGCGAGCTAGCTGCAATGAGCTTCTATATTTTGTCACTTTTTCGTCGTCGTCGG 2340
Db 2610 GCATTCCTGGCGGGAGCGACGATCTTGTATCTTCTCTCGTGTCTTCTGCTTGTGCAATGG 2669
Qy 2341 TACATCAAGGGTCCGGTATGTCCTCTTACCTAATTCCTCACTGGCTGTGTCTT 2400
Db 2670 TATTTGAAGGGTAAAGTGGGTGCCGAGCGGTCTACACTTCTACGGGATGTGGCTCTC 2729
Qy 2401 AGCTTACTGCTCTCAGCATTTGCCCAACAGGCTTATGTTATGACGATCTGTGCAATGGC 2460
Db 2730 CTCTGCTCTCTGTGGCGTGTCCCAAGCGGCGTACGCGTGGACAGAGGTGGCCGCG 2789
Qy 2461 CAGATAGGAGCGGCTCTGCTGTATATGATCACTCTTTTACTCTACCCCCGCGGATTAAG 2520
Db 2790 TCGTGTGGCGGCTGTGTCTCTGTGGGGTGTATGGCGCTGACTCTGTACCAATATTACAG 2849
Qy 2521 ACCCTTCAAGCCGGTTTTTGTGTGTGTGTGTGTACTCTTGTGACCTTGGGGGAAGCTATG 2580
Db 2850 CGCTATATCAGCTGGTGTCTTGTGTGGCTTCAAGTATTTTCTGACAGAGTGAAGCCAA 2909
Qy 2581 GTCCAGAGTGGGACCACTATGAGAGTGGCGGCTGGCCTGTATGGAATATATGAGCC 2640
Db 2910 CTGACGCTGTATTTCCCTCCCTCAACGTCGAGGGGGGCGGACGCGCTATCTTACTC 2969
Qy 2641 GTCCGCAATATTTCAACCAAGTGTGGTGTGTGACATTAACCAAGTGGCTTGGCGGTGTT 2700
Db 2970 ATGTGTGTGTACACCCGACTGTGTATTTTGTACATCAACCAATGTGCTGTGGCTGTTC 3029
Qy 2701 GGGGCTGTACTCTCTTAAGAGTGTCTTGAAGCGCGGCTGTCTTGTGCAAGGCTCAC 2760

Db 3030 GGAACCCCTTGGATTTCTTCAAGCCAGTTTGCTTAAAGTACCTTACTTGTGGCGCTCAA 3089
Qy 2761 GCTTACTAGAGATGTGCAACATTTGGCAAGGCAATCTCGCGGGGGGAGATAGTCCAGATG 2820
Db 3090 GGCCTTCTCCGTTCTGCGGCTTGAAGCGGAAAGATGTATGGAGGCCATTTAGCGCAATG 3149
Qy 2821 GCGCTACTAGCCCTTGGCAGGTGSACTGCACTTACATCTATGACCACTCAACCCCTATG 2880
Db 3150 GTCATATTAAGTTAGGGCGGCTTACTGCACTATGTTTATTAACATCTCACTCTCTT 3209
Qy 2881 TCGGATTTGGGCTGTATAGTGGCTTGGCGGACCTGGCGGCTGGCGGCTTATCATTTTC 2940
Db 3210 CGGGACTGGGCCCAACAACGCTTGGCGAGATCTGGCGGCTGTATGAGCACTGTCTTC 3269
Qy 2941 AGTCCGATGAGAAAGTCACTTGTGTGGGAGCGGAGACAGCTGTGTGGGGAATT 3000
Db 3270 TCCCAATGGAAGCAAGCTCATACGTGGGGGGAGATACCGCGGCTGTGGGTACATC 3329
Qy 3001 TTACACGACTTCCCGTGTCCGCCGACTTGTGTGGAGGTCTCTTGGCCAGCTGAT 3060
Db 3330 ATCAACGCTTGGCTGTCTTCCGCCGCAAGGGCCGGAGATACTGCTCGGGCAAGCGAT 3389
Qy 3061 GGCCTAATCTCCAAGGGGTGAGTCTTCTGCCCATATCACTGTCTTAAGCCCAAGCAGATA 3120
Db 3390 GGAATGTCTCCAAAGGGGTGAGGTTGTGGGCGCCATACAGGCTTACGCCGACAGACA 3449
Qy 3121 CGTGGCTTTTGGGCAACATAGTGTGAGCATGACGAGGCGGAGCAAGAACAGGCT 3180
Db 3450 AGGGGCTCTCTAGGGTGCATTAATACAGCTTATCTGCGGGAGCAAAACCAAGTGGAG 3509
Qy 3181 GGGGAATTCAGGTCTCTGTCCACAGTCACTCACTAGTCTTCTCGGAATCATCTCGGG 3240
Db 3510 GTGAGGTCCAGATTGTGTCACTGTGCTGCCAAACCTTCTGCAACGTGATCAATGG 3569
Qy 3241 GTTTTGTGACTGTCTACATGAGTGTGCAACAAAGATCTGTGGCGGCTTCAAGGGGTCCG 3300
Db 3570 GTGTGTGACTGTCTACACAGGGCGGAGCAAGAGCAATCGGCTCACCAAGGTTCT 3629
Qy 3301 GTCAACGATGTACTCCAGTGTGAGGGGACTTGTATGAGGTGGCCAGCCCTGTGG 3360
Db 3630 GTATCCAGATGTATACATGTATGACCAAGACTTGTGGCTGTGGCTCCGCAAGGT 3689
Qy 3361 ACTAAATCTTGTGAGCGCTGCAAGTGTGAGCGGCTGACCTGTACTGTGACGCGGAAC 3420
Db 3690 AGCGGCTCATTTGACACCTTGCACTTGGCTCCTCGGACTTATCTGTGTACAGAGCAC 3749
Qy 3421 GCTGATGTCACTCCGCTGGAAGACGCGGGGACAAACGGGAGCGTACTTCTCCGAGA 3480
Db 3750 GCGGATGTCAATCCCGTGGCGCGGCGGAGTGAAGAGGGGACAGCTGTGTGCCCCG 3809
Qy 3481 CCTCTTCCACCTTGAAGGGGCTCTCAGAGGCGCGGTGTATGACCCAGGGGCAAGCT 3540
Db 3810 CCGATTTCTACTTGAAGGCTCTCGGGGGGTCTCGCTGTGTGTGCCCCCGGGGACGCG 3869
Qy 3541 GTCGAGTCTTCCGGGCACTGTGTCTCTCGGGGCGTGAAGTATGATTAATTTTCACTC 3600
Db 3870 GTGGGATATTTTATGGGCGCGGCTGTGACCCGAGTGAAGTGAAGCGGTGAATTTATC 3929
Qy 3601 CCGTTGAGACACTGCAATCTGTACGCGGTCCTCCCACTTTATGTAACAACAGACACA 3660
Db 3930 CCTGTGAGAACTTAAGAACAAACATGAGGTCTCCCGGTGTCAAGATTAATCTCTTCCA 3989
Qy 3661 CCTGTGTGCGCCAGACTATCAGTCCGGGTACTTGCATGCCCCGACTGGCAATGGAAG 3720
Db 3990 CCAGTATGTCCCAAGACTTCCAGGTGGCTCACTCTCAATGCTCCACAGGACGGGCAAA 4049
Qy 3721 AGCACCAAGTCTCTGTGCAATGTCTGCTCAGGGGTATTAAGTGTAGTCTTAATCCC 3780
Db 4050 AGCACCAAGTCTCCGGCTGCAATGTGAGCTCAGGGGTATTAAGGTGTCTAGTACTCAACCC 4109
Qy 3781 TCAGTGCTGCAACCTCGGGGTTTTGGGGGTACTTGTCTAAGCAATGCAATCAATCCC 3840

Db 4110 TCTGTGCTGCAACACTGAGGCTTTGGTGTCTTATCATGTCCAAAGGCTCATGGAGTGAATCCT 4169
Qy 3841 AACATTAGAGCTGGAGTCAGAGACTGAGACGCGGGGCGCCATCAGTACTCCACATAT 3900
Db 4170 AACATCAGACCGGGGAGAAACAATTAACACTGGACGCCCCATCAGTACTCCACATAC 4229
Qy 3901 GGGAAATTCCTCGCCGATGGGGGCTGTGCGGGCGGCGCTACGACATCATCATATGTGAT 3960
Db 4230 GGGAGTTCCTTCGCGACGGCGGGTGTCTGGGGGGGCGCTTATGACATATATTTGTGAC 4289
Qy 3961 GAATGCCATCCCTGAGCTTACCAACATCTTGGCATCGGAACATGCTTGTGAAGCA 4020
Db 4290 GAGTGGCACTCCACGAGTCACATCTTGGGCACTCGGCACTGTCTTGAACAAGCA 4349
Qy 4021 GAGACAGCTGGGGTCAACTTAATGTGCTGGCTTACACTAGCCCTCGGGTCAATGACA 4080
Db 4350 GAGATCGGGGGGCGAAGCTGTGTGTCTGACACCGCCACCTCCGGGCTCGTACT 4409
Qy 4081 ACCCCCCACCCAAACATAGAGAGTGGCCCTTGGGAGGAGGCGAGATCCCTTCTAT 4140
Db 4410 GTGCCCCATCCAAACATCGAGAGTGTGCTGTCCACACCGAGAGATCCCTTTTAC 4469
Qy 4141 GGGAGGCGATTCCTCTGTCTTACATCAGGAGGAAGACATGTATCTTGTCCATTA 4200
Db 4470 GGGAAAGGCTATCCCTCTCGAAGTATCAAGGGGGGAGACATCTCATCTTGTCAATTA 4529
Qy 4201 AAGAAAAGTGTACCAAGCTCGGGGCGCCCTTCGGGGTATGGGCTTGAATCAGTGGCA 4260
Db 4530 AAGAAAGAGTCACAGAACTCGCGCAAGCTGGTGCATTTGGGCAATATGCGTGGCC 4589
Qy 4261 TACTACAGAGGGTGTGACGCTCCGTAAATACAACTCAGGAGAGAGTGTGTGCTGCC 4320
Db 4590 TACTACGCGGTCTTGAAGCTGTCCGTATCCGACACAGCGCGCATGTTGTGCTGTGGCA 4649
Qy 4321 ACCGAGCGCTCATGACAGGGTATCTGGGAGCTTTGACTCCGTATGACTGCACTGCAAGTA 4380
Db 4650 ACCGATGCTCATGACCGGCTATACCGGCGACTTGCATCGGAGTATAGACTGCAATACG 4709
Qy 4381 GCGGTACTCAATGTGAGCTTCAGTTAGACCCCAATTCACATCAACACAGAT 4440
Db 4710 TGTGTACCCGACAGTCGATTTCAAGCTTGACCTTACCTTACCATTTGAGAACATACG 4769
Qy 4441 GTCCCTCAAGACGCTGTCTCAGCTAGGCGAGCGCGGGGTGCGACGGGTATGGGAAAGTGT 4500
Db 4770 CTCCCCAGAGATCTGTCTCCGCACTCAACGTCGGGGCAGAGCTGCGAGGGGAGCA 4829
Qy 4501 GGCATTATATAGTATGTTTCCACTGTGAGCGAGCTCAGAGAAATTTTGAACAGTATAGT 4560
Db 4830 GGCATCTACAGATTTTGTGGCAGCGGGGAGCGCCCTCCGGCATGTTGCACTGTCGTC 4889
Qy 4561 CTCTGTGAGCTGCTACAGACGAGGGGCGGATGCTATAGCTTACACCATGTGAGACCAAC 4620
Db 4890 CTCTGTGAGCTGCTACAGACGAGGCTGTGATGAGCTCACGCCCCGAGACTACA 4949
Qy 4621 GTGAGGCTCAGGGCGTATTTCAACAGCGCGGTTTGTGTGCGCAAGACATCTTGAAG 4680
Db 4950 GTTAGGCTACGAGCGTATGAAACACCGGGGCTTCCGTGTCCAGAGCACTTTGA 5009
Qy 4681 TTTTGGAGGCACTTTTACCGGCTTCAACACATAGATGCCACTTCTTTCCAAACA 4740
Db 5010 TTTTGGAGGCGCTTTTACAGGCTTCACTATATAGTCCCACTTTTATCCAGACA 5069
Qy 4741 AAGCAATCGGGGGAATTTTGCATATTAACAGCTTACAGGCTACAGGCTACAGTGTGCTAGG 4800
Db 5070 AAGCAAGTGGGAGAACTTCTTACCTGTAGCGGTACCAAGCCACCGTGTGCGCTAGG 5129
Qy 4801 GCCAAAGCCCCCGCTCTGTGGAGCTCATGTGGAAGTGTGACTCGACTCAAGCCC 4860
Db 5130 GCTCAAGCCCTTCCCATGTGTGGACCAATGTGGAAGTGTGATTCGCTCAAGCCC 5189
Qy 4861 ACACTGTGGGCCCCCACTTCTCTGTACCGCTTGGGCTCTGTATCAACAGAGTCAAC 4920
Db 5190 ACCCTCATGGGCAACACCCCTGTATATAGACTGGGCGCTGTTCAGAAATGAATCAC 5249

Qy 4921 CTCACATCCGCTGACGAATATCATGCCACCGTCAATGCAAGCCGACCTTGAAGTCAATG 4980
Db 5250 CTGACGACCCAGTCAACAAATATCATATATGATGATGCGGCGACCTGGAGGTGTC 5309
Qy 4981 ACCAGACATGGGTCTTGGCAGGGGAGTGTGGCGGCGCTGCGGCTATTCCTGGCG 5040
Db 5310 ACAGACACCTGGGTGCTGTGGCGGGTCTGTGGCTTGTGGCGGGTATTCGCTGCA 5369
Qy 5041 ACCGGGTGTGTGATCATGCGCCGCTTGCACATTAACAGCGAGCGTGTGGCGCG 5100
Db 5370 ACAGGCTGCTGTGATAGTGGCAGGGTGTGTGCTTGTCCGGAAAGCGGCAATCATACT 5429
Qy 5101 GACAGAGGTCCTCTATAGAGCTTTGATGAGATGAGGAATGTGCTTATGGCGGGCT 5160
Db 5430 GACAGGAAGTCTCTACCGAGAGTTCGATGAGATGGAAGAGTGTCTTACACTTACCG 5489
Qy 5161 CTCATTGAAGAGGGGAGCGGATAGCGGAGATCTGAAGTCCAAAGCTTATTTG 5220
Db 5490 TACATGAGCAAGGAGTATGATGCTCGCGAGCAATTAAGAGAAAGGCTTGGGCTCTG 5549
Qy 5221 CAGCAAGCTTCCAAACAGTCAAGCATTAACCACTGTGAGGCTTATGGCCAG 5280
Db 5550 CAGACGCGCTCCGTCAGGCGAGGTTATCGCCCTGTGTCCAGACCAACTGGCAAAAA 5609
Qy 5281 GTAGAACATTTCTGGGCGAAACATGTGGAACTTGAAGGGGCAATCCAAATCTCGCA 5340
Db 5610 CTGAGACCTTCTGGGGGAAAGCATATGTGAACTTCACTAGTGGAAACAATCTTGGCG 5669
Qy 5341 GCACTATCAACACTGCGCAGGGAACCTGCAAGTCTTCCATGATGCGGCTCAAGTCCGCC 5400
Db 5670 GCTTGTCAACGCTGTGCTGTGTAACCGGCATGCTTCAATGATGCTTTTACAGCTGT 5729
Qy 5401 CTCACAGTTCGCTGTCAACAGACCAACTATCTTCTCAACATTTTGGGGGCTGGCTA 5460
Db 5730 GTCAACAGCCCACTTAACCACTTACCAAACTCTCTTCAACATATTTGGGGGGTGGTG 5789
Qy 5461 GCATCCCAATTGCAACACCGCGGGGGCCACAGGCTGTGTCTCAAGTGGCTAGTGGGA 5520
Db 5790 GCTGCCAGTCCCGCGCCCGGCGGTGCGTACTGCTTGTGGCGCGTGTAGCTGGC 5849
Qy 5521 GTCGCGTAGGAGCATATAGGCTTAGGTAAAGTGTAGTGAACATCTGCGAGGATAGT 5580
Db 5850 GCGGCATGCGGAGTGTGTGGAATGCGGGAAGTCTCATAGACATCTTGTGAGGGTATGCG 5909
Qy 5581 GCGGCAATTTGGGGGCTCTGTGCAATTCAGATCATGTCTGCGAGAAACCTTCATG 5640
Db 5910 GCGGGCGTGGCGGAGCTTGTGCAATTCAGATCATGAGCGGTAGGTCCTCCACG 5969
Qy 5641 GAGGATGTGTCAACTTGTGCTGCTGGAATTCGTCTCCGGGTGCTTGTGATGGAGTGC 5700
Db 5970 GAGGACCTGTCAATCTACTGCTCGGCATCTTGTGCGGAGCCCTCGATGTGGGCTG 6029
Qy 5701 ATCTGCGGCGCAATTCGTGCGCGACAGTGGAGCCGGGGGAAGGCGCTGCAATGATG 5760
Db 6030 GTCTGTGAGCAATATCTGCGCGGCAAGTTGGCTCGGGGAGAGGGGCAAGTGTGATG 6089
Qy 5761 AATGACTATTCCTTTGTCTTCAAGAGAAATCAGTGTGCGCCCACTACGTAAGTACG 5820
Db 6090 AACGGCTGATACCTTTCGCTCCCGGGGAAACAGTGTTCCTCCCACTACGTCGCG 6149
Qy 5821 GAGTGGATGCGTGGAGGCTGAGCCCAACTACTTGGCTCCCTTACCAATACAGCTG 5880
Db 6150 GAGAGGATGCAAGTCTCGCGCTGCTGCAATCTCAGAGGCTTCACTGTAACTCAAGCTC 6209
Qy 5881 CTCAGAAAGCTCCAAACTGATATCTAGGAGTGTGCCCATCCATGCGGCGGCTGTG 5940
Db 6210 CTGAGGAGACTGCAACAGTGAATACCTGGAAGTGTACCACTCAAGCTCCGGTTCTGG 6269
Qy 5941 CTCGCGATGTGTGGAGCTGGGTTTGCACCACTTCAACAGCTTTAAAAATTTGCTGACC 6000
Db 6270 CTAAAGGAGATCTGAGGACTGAGTATGCGAGGTGTGAGGACTTTTAAAGCTTGGCTAAA 6329

QY 6001 TCNAATTATTTCCCAAGATCCCGCCTCCCTTTGTCTCTGTCAAAAGGGTACAG 6060
Db 6330 GCTAAGCTCATGTCCACAGCTGCTGGATATCCCTTTGTGTCTCTCCAGCGGGTATTAAG 6389
QY 6061 GGGCTGTGGGCGGCACTGGCATCTATACACAGGTGTCTTGGCGGCAATATCTCT 6120
Db 6390 GGGGTCTGGCGAGTGGACGGGATCATGCACTGGCTGCACTGTGGAGCTGAATCACT 6449
QY 6121 GGCAATGTCCGCTTGGGCTCCATGAGATCACGGGGCTTAAGACTGCAATATCTGG 6180
Db 6450 GGACATGTCAAAAACGGGACGATGAGATCTGTGGTCTTAGAGACTGAGAAATATGG 6509
QY 6181 CAGGGGACCTTCTCTATCAATTTGTTACACGAGGGCCAGTGGCTGGGAAACCGCGCA 6240
Db 6510 AGTGGACCTTCCCATTTAATGCTACACAGGGGCCCTGTACCCCTCTCTGGCCG 6569
QY 6241 AACTTAAAGTTCGCATCTGGAGGGTGGCGGCTTCAAGTACGGGAGGTACCGACAC 6300
Db 6570 AACTACAGTTTGGCTATGAGAGGGTGTCTGCAAGAAATGTGAGATTAAGGCAAGT 6629
QY 6301 GGGTCATACCACTACATAACAGSACTCACACTGATACTTGAAGTCCCTGCAACTA 6360
Db 6630 GGGGACTTCCACTACGTACGGGTATGACTACTGACAACTCAATGCGCGTGCAGAGTC 6689
QY 6361 CCTCTCCGAGTTCTTTTCTGGGTGGAAGAGTGCAGATCAATAGTTTGGCCCA 6420
Db 6690 CCATGCCCCGAATTTTTCACAGAAATGGAGCGGGGTGCGCTTACATAGGTTTGGCCCC 6749
QY 6421 CCGAAGCGTTTTCGGGATGAGGTCTGTCTGCGTGGCTTAATTAATTTGTGTGTC 6480
Db 6750 TCGAAGCCCTTGTGCGGAGAGGATATCATTCAGATGAGATCTCAAGAAATACCGGTA 6809
QY 6481 GGGTCCCAAGTCTCTTGGCAGCCTTGAACCGCACACAGAGTATGATGTCCATGTA 6540
Db 6810 GGGTGCATTTACTCTTGGAGCGCGAACGGACGGTGGCTGTGAGTCACTCACTCACT 6869
QY 6541 GATTCATCTCATATACAGCGGAGACTGCAGCGGGGCTTTAGCGGGGGTCAACCCCA 6600
Db 6870 GATCCCTCCATATTAACAGAGAGGGCGCGGCGCAAGTTGGGAGGGATCAACCCCC 6929
QY 6601 TCCGAGGCAAGCTCTCGGCGAGCGAGTATTCGAGACCATGCTGCCAGCCACTGAC 6660
Db 6930 TCTGTGGCCAGCTCTCGGCTAGCCAGTATCTCGTCACTCTCAAGGCAACTTGCACC 6989
QY 6661 ACCCAGGCAAGCCTATGATGTGACATGTGTGATGCTAACT-----GTTCT 6708
Db 6990 GCTAACATGACTCCCTGATGTGAGTCAATAGAGGCCAATCTCTATGAGGCAAGAG 7049
QY 6709 ATGGGGGGGATGTGATCTCGATAGAGTGTGGTCCAAAGTGTCTTGGACTCTCTC 7068
Db 7050 ATGGGCGGCAATCATCAAGGATGTGATGAGTCAAGAAACAAAGTGTGATTTCTGGACTCTTC 7109
QY 6769 GACCAATGTGTGAAGAAAGAGGAGCCTTGAGCCTTGATACCATGAGATACATGTCTC 6828
Db 7110 GATTCGCTTGTGGCGAGAGAGGAGCGAGGAGATCTCGTACCGGCAAAATCTGTGG 7169
QY 6829 CCCAAGAGGTTCCCAACAGCTTACCGGCTTGGGCAAGGCGCTGATTAACAACCG 6888
Db 7110 AAGTCTCGAGATTTGGCCAGGCGCTGGCCGTTTGGGGCGGCGGAGATTAACCCCG 7229
QY 6889 CTTTGTGAATCTGTGAAGAAAGCCAGATTACCAACCGGCACTGTTGGGGGCTGTCTCTC 6948
Db 7230 CTAGTGGAGACGTGAGAAAGCCGACTACGAACCACTGTGTTCATAGCTGTCTCGCTT 7289
QY 6949 CCTCTCTAGGAAACCCCGAGCGCTCCCAAGAGGCGCGGAGCATGGGCTTAAGT 7008
Db 7290 CCACCTTCAAAAGTCCCTCTGTGTCTCGGCTTGGAAAGAGCGAGTGTGTCTCACT 7349
QY 7009 GAGGACTCATGAGATGCGCTTCAAGAGCTGCGCATTTAAGTCTTTTGGCGAGCCGCC 7068
Db 7350 GAATGAACCTTATCTAGCTGCTTGGCGAGCTGCGCACCAAGAGTTTGGGAGCTCTCA 7409
QY 7069 CCAAGCGCGATTCAAGGCTTTTCAAGGGGCGGGCGCTGCCGATTTCCGGAGTCAAGC 7128

Db 7410 ACTTCCGGATTACGGGCGACAAATACGACAAATCTCTGAGCCGCC----- 7457
QY 7129 CTTCCATGATGAGTGTGGCCCTTTTCGAGACAGGTTCCATCTTTCAATGCCCCCTCGAG 7188
Db 7458 CTTTCTGGCTGCCCCCGACCTCCAGGCTGTAGTCTTATCTTCCATGCCCCCTGGAG 7517
QY 7189 GGGGAGCTTGTGAGATTCAGACTGTGAGCCTGAGCAGGTAGAGCCCAACCCGCCAG 7248
Db 7518 GGGGAGCTGTGGGATTCGGA----- 7537
QY 7249 GGGGGGGTGGCAGCTCCCGCTCGACTGGGGTCTGTGTCTATCTTGTCTCGAGAGAGC 7308
Db 7538 -----TCTTACGACAGGGTCAATGATCAACGTCAGTATGTAGAGCAACGGG 7583
QY 7309 GACTCGGTGTGTGCTGCTCAATGTCAATCTCTGAGACGGGGCTTAAATCTCTGT 7368
Db 7584 GAGGATGTGTGTGCTGCTCAATGTCTTATCTTGTGACAGGGGCACTGTCAACCCGTGC 7643
QY 7369 AGTCCGAAGAGAGAGATTACGATTAACCCCTTGAAGCACTCCCTGTGGGATATCAC 7428
Db 7644 GCGCGGAAGAACAGAAACCTGCCATCAATGACATTAAGCACTCGTTGCTACGTCACAC 7703
QY 7429 AACAGGTGTCTGTACCAACAAAGAGCGCTCACTAAGGCTTAAAGGTAATTTT 7488
Db 7704 AATTGTGTATTCACACACCTCACGAGTGTCTTGCAAAGGCAAGAAAGTCACTTT 7763
QY 7489 GATAGGATGCAAGTGTCTGACTCTTACTGACTGACTGCTTAAAGGACATTAAGCTAGC 7548
Db 7764 GACAGACTGCAAGTTTGTGACAGCCATTAACAGAGCTACTCAAGAGGTTAAAGCAGC 7823
QY 7549 GCCTCCAGAGTCAACCGAAGGCTCTCACCATGAGAGAGCTTGCCAGTTAACCCACCC 7608
Db 7824 GCGTCAAAAGTGAAGGCTTAACCTTGTATCGTATAGAGAGCTTGACGCTGACGCCCA 7883
QY 7609 CATTTGCAAGATCTAAATATGGGTTTGGGGCTTAAGAGGTCCGCACTTGTCCGGAG 7668
Db 7884 CACTCAGCCAAATCCAAAGTTTGTATGAGGCAAAAGAGTGTCTGTCATGCCAAGAA 7943
QY 7669 GCGTTAACCACTCAAGTCCGCTGAGAGAGCCTCTGAGAGACTCAGAAACACCAATT 7728
Db 7944 GCGTTAACCACTCAAGTCCGCTGAGAGAGCCTTCTGAGAGCAATGTTAACCAATTA 8003
QY 7729 CCCAACAATTAATGSCCAAAATGAGGTGTTCTGCTGAGACCCACAAAGGGGCAAG 7788
Db 8004 GACACTACATATGCTTAAGAACGAGGTTTCTGTGCTCAAGCTGAGAAAGGGGCTCGT 8063
QY 7789 AAAGCAGTCTGCTTATCTGTTTACCTGACCTTGGGCTGAGGCTGTGCGAAGAAATGGC 7848
Db 8064 AAGCCAGCTGTCTCATGTGTGCTCCGATCTGGGGGTGCGGCTGTGCAAAAGATGGCT 8123
QY 7849 CTTTATGACATTAACAATAACTTCTCAGGCGGATGAGGGGCTTATGATTTCCAG 7908
Db 8124 TTGTATGAGCTGTGTTACAAAGCTCCCTTGGCCGTATAGGAAAGCTCTTACGATTTCCA 8183
QY 7909 TATTTCCCGCTCAGCGGGTGAAGTTTCTTGAAGGATGAGCGGAAAGAGACCT 7968
Db 8184 TACTCACAGGACAGCGGGTGAATTTCTGTGCAAGCTGAGAGTCCAAAGAAACCCCA 8243
QY 7969 ATGGGTTTTCTGTATATACCGGATGCTTTGACTCAACGTCATGAGAGAGACATCAG 8028
Db 8244 ATGGGTTTTCTGTATATACCGGCTGCTTGTACTCAAGTCACTGAGAGGCAATCCGT 8303
QY 8029 ACTGAGAGTCCATATATCGGGGCTGCTCTTTCAGGAGGAGGCGACACTGCAATAC 8088
Db 8304 ACAGAGAGGCAATCTACAAATGTGTGACCTGACCCCAAGCCCGCTGCGCATCAAG 8363
QY 8089 TCGCTAACTGAGAGCTTTACGTGAGAGGCGCTATGTTCAACGAGAGGCAACCTGCG 8148
Db 8364 TCCCTCACGAGAGGCTTTATGTTGGGGGCGCTTACCAATTCAAGGGGGAGAACTCG 8423
QY 8149 GGGTACAGGCTTGGCGCGCAGCGGGGTGCTACCACTAGCATGAGGAAACATATCA 8208

Db 8424 GGCTATGCGAGGTGCGCGGAGCGGCTACTGACAACTAGCTGTGTAAACCCCTCACT 8483
QY 8209 TGTACGTGAAGGCTTAGCGGCTTTGTAAGCTGAGGGATATCGGCCCAATGCTG 8268
Db 8484 TGTACATCAAGGCGCGGCGAGCTGTGAGCCGAGGGCTTCAGAGCTGCACCAATGCTC 8543
QY 8269 GTATGCGGCGATGACTTGTGTGATCTCAGAAAAGCCAGGGAGCCGAGAGGAGCG 8328
Db 8544 GTGTGCGGCGAGCTTAGTGTATCTGTGAAGCGCGGGGTTCAGAGAGGAGCGCGCG 8603
QY 8329 AACCTGAGGCTTTCAGCGAGGCTATGACCAAGTATTTCTGCCCTCTGTGACCCCCC 8388
Db 8604 AGCTGTAGAGCTTTCACGAGGCTATGACCAAGTATTTCTGCCCTCTGTGACCCCCC 8663
QY 8389 AGACCGAGTATGATCTGAGGCTATGACCAAGTATTTCTGCCCTCTGTGACCGCTG 8448
Db 8664 CAAACGAAATAGACTTGTAGACTCATATCATATGCTCTCAAGTGTCAGTGCGCCAC 8723
QY 8449 GGCCCAAGGCGCGCGCAATACTACTGACAGAGACCTTACCACTCCATCGCCCG 8508
Db 8724 GACGGCGCTGAAAGAGGCTTACTACTCACCCTGACCTTACACCCCTCGCGAGA 8783
QY 8509 GCTGCTGGGAAACAGTTAGCACTCCCTGTCAATTCATGAGCTGGGAAACATCATCG 8568
Db 8784 GCTGCTGGGAGAGAGCAAGACATCTCAATTCCTGTAGGCAATATCATG 8843
QY 8569 TAGGCCCCGACATATGGGCTCGATGCTCTGATGACACACTTCTTCATTTCTCATG 8628
Db 8844 TTTGCCCCCAGCTGTGGCGAGATATGATGATGACCAATTTCTTTAGGCTCTTATA 8903
QY 8629 GCTCAAGACAGCTGAGACCAAGAACCTCACTTTGAGATGTAGAGGCGGTGATCTCCG 8688
Db 8904 GCCAGGAGCAAGCTTGAACAGGCGCTGTGATGCGAGATCTAGGCGCTGTACTCATTA 8963
QY 8689 AGTCCCTTGAAGCTCCAGCTATATTAATTGAAGGTTACATGGGCTTGAAGCTTTTCTG 8748
Db 8964 GAACCACTTGTATTAATCTCCATCATTTCAAGACTGCGCTCAGCGCATTTTCACTC 9023
QY 8749 CACGATATCACTCCCAAGAACTGACACAGGCTGTGAGCTTCAGGCTTCAGAAAATTGGGGCG 8808
Db 9024 CACGTTACTCTCAGGTGAATTAATAGGCTGGCGGATGCTCAGAAAATTGGGGTA 9083
QY 8809 CCAACCTCTCAGAGGCTGGAAGAGCGGAGCAAGTCAAGGCGGCTCCCTCATCTCCGT 8868
Db 9084 CCGCCCTTGGAGCTTGAAGACACCGGCGCGAGCGTCCGCGCTTAGGCTTCTGGCCAGA 9143
QY 8869 GGGGGGAGAGGCGGCTTGGCGTGAATCTTCAATTGGGCGGTGAAGACCAAGCTC 8928
Db 9144 GGAAGGAGGGGCTGATATGGAAGTACTTTCAACTGGGCAGTAAAGCAAGCTC 9203
QY 8929 AAATCTACTCTCATTTGCGGAAGCGCGCTCTGGAATTTATCAGCTGTTCAACCTGCGC 8988
Db 9204 AAATCTACTCTCAATAGCGGCGCTGCGAGCTGGAATTTGCGGCTGTTCACCGCTGGC 9263
QY 8989 GCCGCGGGGCGCAATTTATCAGAGGTGTGCGGTGCCGAGACCGGCTTATGCTCTT 9048
Db 9264 TACAGCGGGGAGCAATTTATCAGAGGTGTCTCATGCGGCGCGCTGAGCTGCTGTT 9323
QY 9049 GGCCTACTCTACTTTTGTAGGGGTAGGCTTTTCTACTCCCGCTCG 9098
Db 9324 TGCTTACTCTGCTGTGCGAGGGGTAGGCTTACTCTCTCCCAACCG 9373

Search completed: November 1, 2005, 04:24:43
Job time : 1140 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2005, 15:35:06 : Search time 285 Seconds
(without alignments)
4115.945 Million cell updates/sec

Title: US-09-980-559-2

Perfect score: 16120

Sequence: 1 MSTNPKPQKTKRNTNRBPQ.....LFGLLLPVGVGLFLLPAR 3033

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: geneseqp1980s:***
2: geneseqp1990s:***
3: geneseqp2000s:***
4: geneseqp2001s:***
5: geneseqp2002s:***
6: geneseqp2003as:***
7: geneseqp2003bs:***
8: geneseqp2004s:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	16120	100.0	3033	4	AAB59172	Aab59172 Protein e
2	16120	100.0	3033	4	AAB31168	Aab31168 Amino aci
3	16120	100.0	3033	4	AAB30729	Aab30729 Amino aci
4	15824	98.2	3033	2	AAB33538	Aab33538 NANBH vir
5	15645	97.1	2940	7	ADCA2917	Adca2917 Hepatitis
6	14943	92.7	3033	5	ABG30688	Abg30688 Human HCV
7	13943	86.5	3033	2	AAR33539	Aar33539 NANBH vir
8	13805	85.6	3033	2	AAR33214	Aar33214 NANBH vir
9	13049	80.9	3015	4	AAB30730	Aab30730 Amino aci
10	13049	80.9	3015	4	AAB30732	Aab30732 Amino aci
11	12925	80.2	3015	4	AAB30733	Aab30733 Amino aci
12	12925	80.2	3015	4	AAB30731	Aab30731 Amino aci
13	12925	74.7	3010	5	ABG32459	Abg32459 Hepatitis
14	12033	74.6	3010	5	ABG32455	Abg32455 Hepatitis
15	12033	74.6	3010	5	ABG32451	Abg32451 Hepatitis
16	12028	74.6	3010	5	ABG32453	Abg32453 Hepatitis
17	12027	74.6	3010	5	ABG32460	Abg32460 Hepatitis
18	12027	74.6	3010	5	ABG32454	Abg32454 Hepatitis
19	12026	74.6	3010	5	ABG32461	Abg32461 Hepatitis
20	12025	74.6	3010	5	ABG32458	Abg32458 Hepatitis
21	12025	74.6	3010	5	ABG32457	Abg32457 Hepatitis
22	12025	74.6	3010	5	ABG32452	Abg32452 Hepatitis
23	12022	74.6	3011	5	ABG32456	Abg32456 Hepatitis
24	11998	74.4	3010	7	ADP88597	Adp88597 Hepatitis
25	11990	74.4	3010	2	AAR68864	Aar68864 Hepatitis

26	11989	74.4	3010	2	AAY06423	Aay06423 Non-A, no
27	11986	74.4	3010	2	AAR82694	Aar82694 Partial H
28	11984	74.3	3010	2	AAR20111	Aar20111 Non-A, no
29	11984	74.3	3010	2	AAR20091	Aar20091 Non-A, no
30	11977	74.3	3011	5	AAU84597	Aau84597 HCV poly
31	11973	74.3	3011	2	AAW77397	Aaw77397 Hepatitis
32	11973	74.3	3011	6	ABP71460	Abp71460 Amino aci
33	11973	74.3	3011	8	ADH79949	Adh79949 E2 HCV en
34	11973	74.3	3011	8	ADL72983	Adl72983 Hepatitis
35	11973	74.3	3012	5	AAU99289	Aau99289 Hepatitis
36	11973	74.3	3012	6	ABU61848	Abu61848 HCV H7 c
37	11966	74.2	3011	4	AAB31169	Aab31169 Amino aci
38	11966	74.2	3011	5	AAU79221	Aau79221 Hepatitis
39	11966	74.2	3011	8	ADJ56744	Adj56744 Hepatitis
40	11966	74.2	3011	8	ADJ64256	Adj64256 Hepatitis
41	11966	74.2	3180	6	ABG73195	Abg73195 MKO-Z vir
42	11964	74.2	3011	2	AAW98020	Aaw98020 Infectiou
43	11963	74.2	3010	2	AAR68622	Aar68622 HCV prote
44	11958	74.2	3010	2	AAW98022	Aaw98022 Infectiou
45	11958	74.2	3010	4	AAB31170	Aab31170 Amino aci

ALIGNMENTS

RESULT 1

AAB59172
ID AAB59172 standard; protein; 3033 AA.

AC AAB59172;

DT 21-MAR-2001 (first entry)

DE Protein encoded by infectious Hepatitis C virus 2a genotype.

KW GBV-B; hepatitis C virus; HCV; vaccine.

XX Hepatitis C virus.

XX WO200075337-A1.

XX 14-DEC-2000.

XX 02-JUN-2000; 2000WO-US015293.

XX 04-JUN-1999; 99US-0137694P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Bukh J, Yanagi M, Emerson SU, Purcell RH;

XX WPI; 2001-091214/10.

PT New infectious nucleic acids of the GB virus-B clone, useful for indirectly studying the molecular properties of hepatitis C virus (HCV) and in developing vaccines and therapeutics for HCV.

XX Disclosure; Page 82-94; 96pp; English.

XX The present invention relates to GB virus-B. The nucleic acid molecules of the invention are useful for indirectly studying the molecular properties of hepatitis C virus (HCV). The infectious nucleic acid sequence of the GB virus-B clone and the HCV/GBV-B chimeras may be used in the development of vaccines and therapeutics for HCV

XX Sequence 3033 AA;

XX Query Match 100.0%; Score 16120; DB 4; Length 3033;

XX Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;

XX Matches 3033; Conservative 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRBPQGVGVGLFLLPARGLGVRAKTRKTSRSQPRG 60

Db 1 MSTNPKFORKTKRNTNRNRPDVKFPGGQIVGVYILLPRRGPRLGVARATKRTISERSQPRG 60
 Qy 61 RROP1PDRRSTGSKMGKPGYPMPLYNENELGNAWMLSPRGRSPSMGPNDPHRSRNVG 120
 Db 61 RROP1PDRRSTGSKMGKPGYPMPLYNENELGNAWMLSPRGRSPSMGPNDPHRSRNVG 120
 Qy 121 KVIDTLTCGPADLMGYIPVVGAPLGGVARRALAHGVRVLEDCVNPATGNLPQCSFSLPFLA 180
 Db 121 KVIDTLTCGPADLMGYIPVVGAPLGGVARRALAHGVRVLEDCVNPATGNLPQCSFSLPFLA 180
 Qy 181 LLSGITTVPAAEYKXNSTGYMTNDCNTNSITMQLAAVLAHPCVPCCKVGNASQCM 240
 Db 181 LLSGITTVPAAEYKXNSTGYMTNDCNTNSITMQLAAVLAHPCVPCCKVGNASQCM 240
 Qy 241 PVSBNVAVORPGALTOGLRTHIDMVVMSATLCSALYVGDLCGGVMLAAQMFIVSPQHWF 300
 Db 241 PVSBNVAVORPGALTOGLRTHIDMVVMSATLCSALYVGDLCGGVMLAAQMFIVSPQHWF 300
 Qy 301 VQDCNCSITPRTTIGHMAMDMMNNSPTATMLAYMRVBEVIDIISGAHWGVFGLA 360
 Db 301 VQDCNCSITPRTTIGHMAMDMMNNSPTATMLAYMRVBEVIDIISGAHWGVFGLA 360
 Qy 361 YFSMOGMAKVVVLLLAAGVDARTHTVGSAAQTGRLTSLFDMGPRKIQLVNTNGSW 420
 Db 361 YFSMOGMAKVVVLLLAAGVDARTHTVGSAAQTGRLTSLFDMGPRKIQLVNTNGSW 420
 Qy 421 HINRTALNCNDSLHTGFIASLFYTHSFNSGCCPERMSACRSIEAFRVGMALQYEDNVN 480
 Db 421 HINRTALNCNDSLHTGFIASLFYTHSFNSGCCPERMSACRSIEAFRVGMALQYEDNVN 480
 Qy 481 PEDMRPYCMHYPPROCGVVSATKVCSPVYCTPSPVVVGTTRIGAPTYTWGENEDVFL 540
 Db 481 PEDMRPYCMHYPPROCGVVSATKVCSPVYCTPSPVVVGTTRIGAPTYTWGENEDVFL 540
 Qy 541 LNSRPRPLGSMFGCTMNNSSGYTKTCGAPCRPRADFNASTDILCPDRCRKHDPDTYTK 600
 Db 541 LNSRPRPLGSMFGCTMNNSSGYTKTCGAPCRPRADFNASTDILCPDRCRKHDPDTYTK 600
 Qy 601 CGSGPMLTPRCLIDYPYRLMHYPTCVNYTIFKIRMYGVGVHRLTAACNTRGRDCLAD 660
 Db 601 CGSGPMLTPRCLIDYPYRLMHYPTCVNYTIFKIRMYGVGVHRLTAACNTRGRDCLAD 660
 Qy 661 RDRSOLSPHLSTTBEMAILPCSYSDLPALSTGLIHHONTVDVQPMYGLSPALTKYIVRW 720
 Db 661 RDRSOLSPHLSTTBEMAILPCSYSDLPALSTGLIHHONTVDVQPMYGLSPALTKYIVRW 720
 Qy 721 BMYLLFLULADARVACIMWMLILGOABALEKVLILHSAASCMGFYFYIFPAAW 780
 Db 721 BMYLLFLULADARVACIMWMLILGOABALEKVLILHSAASCMGFYFYIFPAAW 780
 Qy 781 YIKGRVVPATYSLTGLMSFSLILLALPOQAYAYDASVHGOIGALLVMTLFTLPGYK 840
 Db 781 YIKGRVVPATYSLTGLMSFSLILLALPOQAYAYDASVHGOIGALLVMTLFTLPGYK 840
 Qy 841 TILSRFLMWLCYLLTLGBAMVQEMAPMOVRGCGDGIIMAVAFYPCGVFDTIKMLIAVL 900
 Db 841 TILSRFLMWLCYLLTLGBAMVQEMAPMOVRGCGDGIIMAVAFYPCGVFDTIKMLIAVL 900
 Qy 901 GPAYLLGALTRVYFVRAHALLRMCTMARHLAGRVVOMALLAGRWGTGYIYDHLTPM 960
 Db 901 GPAYLLGALTRVYFVRAHALLRMCTMARHLAGRVVOMALLAGRWGTGYIYDHLTPM 960
 Qy 961 SDMAASGLRDIAVAVEPIIFSPMEKXIVWGAETPAACGDIILHGLPVSARLGREVLLGPAD 1020
 Db 961 SDMAASGLRDIAVAVEPIIFSPMEKXIVWGAETPAACGDIILHGLPVSARLGREVLLGPAD 1020
 Qy 1021 GYTSKGSLSLAPITAYAOQTGGLIGTYVSWTGBDKTBQGEIOVISTVQSTFGTISG 1080
 Db 1021 GYTSKGSLSLAPITAYAOQTGGLIGTYVSWTGBDKTBQGEIOVISTVQSTFGTISG 1080
 Qy 1081 VLMYVYVAGAKTLAGRGPVTOMVYSABEDLVGMPSPGCTKSLEPCTCGAVDLVLTNRN 1140
 Db 1081 VLMYVYVAGAKTLAGRGPVTOMVYSABEDLVGMPSPGCTKSLEPCTCGAVDLVLTNRN 1140

Qy 1141 ADVIPARRRGDKRGALLSPRPLSTLKSGSGGVLCPRGHAVGVFPAAVCSRVAKSIDFI 1200
 Db 1141 ADVIPARRRGDKRGALLSPRPLSTLKSGSGGVLCPRGHAVGVFPAAVCSRVAKSIDFI 1200
 Qy 1201 PVEITLDIVTRSPFSDNSTPPAVPOTVOGYLHAATGSGSKTKVVAAYAAQGYKVLVNP 1260
 Db 1201 PVEITLDIVTRSPFSDNSTPPAVPOTVOGYLHAATGSGSKTKVVAAYAAQGYKVLVNP 1260
 Qy 1261 SVAATLPGAYLSKXAGINPNIRTVGRVTTGAPITTSYGVFLADGGCAGAYDIIICD 1320
 Db 1261 SVAATLPGAYLSKXAGINPNIRTVGRVTTGAPITTSYGVFLADGGCAGAYDIIICD 1320
 Qy 1321 ECHAVDSTIIIGITVLDOAETAGVRLTVLATPSPGVTTPHPNIEVALQOGBEIPFY 1380
 Db 1321 ECHAVDSTIIIGITVLDOAETAGVRLTVLATPSPGVTTPHPNIEVALQOGBEIPFY 1380
 Qy 1381 GRAIPLSYIKGGRHLIFCHSKKKCBELAALLRGMGLNSVAYYRGJLDVSVIPQGDVVVVA 1440
 Db 1381 GRAIPLSYIKGGRHLIFCHSKKKCBELAALLRGMGLNSVAYYRGJLDVSVIPQGDVVVVA 1440
 Qy 1441 TDALMTGYTGDPDSYIDCNVAVTQVDFSLDPTFITTOIVPQDAVSRQRGRGTGRGL 1500
 Db 1441 TDALMTGYTGDPDSYIDCNVAVTQVDFSLDPTFITTOIVPQDAVSRQRGRGTGRGL 1500
 Qy 1501 GIYRVYSTGERASGMFDSVVLCECYDAGAWYELTPESETTVLRAYFNTPGLPVCODHLE 1560
 Db 1501 GIYRVYSTGERASGMFDSVVLCECYDAGAWYELTPESETTVLRAYFNTPGLPVCODHLE 1560
 Qy 1561 FWEAVFTGITHIDAFILSOTKQSGENFAYLTAYQATVCARAKAPPPSMDVMKCLTRLP 1620
 Db 1561 FWEAVFTGITHIDAFILSOTKQSGENFAYLTAYQATVCARAKAPPPSMDVMKCLTRLP 1620
 Qy 1621 TLVGPPLLYRGSVTNEVTLTHPTKYATACQADLEVMSTWLAGVLAAYACIA 1680
 Db 1621 TLVGPPLLYRGSVTNEVTLTHPTKYATACQADLEVMSTWLAGVLAAYACIA 1680
 Qy 1681 TGCVCILGRHINORAVVAPDXEVLVEAFDEMEBCASRAALIEBQRIAEMLSKIOGL 1740
 Db 1681 TGCVCILGRHINORAVVAPDXEVLVEAFDEMEBCASRAALIEBQRIAEMLSKIOGL 1740
 Qy 1741 QOASKOADIOPTVOASWPKVQFPAKHMNNTISGITYLAGISTLPQNPVAVSMAAFSA 1800
 Db 1741 QOASKOADIOPTVOASWPKVQFPAKHMNNTISGITYLAGISTLPQNPVAVSMAAFSA 1800
 Qy 1801 LTPSLSTSTIILNLTGMLASQIAPAGATFVVVSGLGAUVGSIIGLKVVDLILAGY 1860
 Db 1801 LTPSLSTSTIILNLTGMLASQIAPAGATFVVVSGLGAUVGSIIGLKVVDLILAGY 1860
 Qy 1861 AGISGALVAFKIMSGEKPSMEDVNLPGIISPGALVWGVICAILRRHVGBGAVQNM 1920
 Db 1861 AGISGALVAFKIMSGEKPSMEDVNLPGIISPGALVWGVICAILRRHVGBGAVQNM 1920
 Qy 1921 NRIAPASRGHVAAPHYVTESDASQRYQLLSITLSLRLRNWITEDPICGGSW 1980
 Db 1921 NRIAPASRGHVAAPHYVTESDASQRYQLLSITLSLRLRNWITEDPICGGSW 1980
 Qy 1981 LRDVMDVCTIITDFGNWLTSLPMPGLPFVSCOKGKGVWAGTGMTTRPCPGANIS 2040
 Db 1981 LRDVMDVCTIITDFGNWLTSLPMPGLPFVSCOKGKGVWAGTGMTTRPCPGANIS 2040
 Qy 2041 GNVRIGSMRITGPKTCMNIWQSTFPCINCYTBQCVKPAKPNKVAIMRWAASEYAEVTOH 2100
 Db 2041 GNVRIGSMRITGPKTCMNIWQSTFPCINCYTBQCVKPAKPNKVAIMRWAASEYAEVTOH 2100
 Qy 2101 GSYHYITGLTIDNLKVPQCLBPBPFSWVDGQIHRFAFPTPRFRDEVSFCVGLNSFV 2160
 Db 2101 GSYHYITGLTIDNLKVPQCLBPBPFSWVDGQIHRFAFPTPRFRDEVSFCVGLNSFV 2160
 Qy 2161 GSOLPCDPRPDVVMMSMLTDPSSHITAEFAARLARSPSEASSASQSLAPSLATCT 2220
 Db 2161 GSOLPCDPRPDVVMMSMLTDPSSHITAEFAARLARSPSEASSASQSLAPSLATCT 2220

QY 2221 THGKAYDVDMVANLFWGSDVTRIESGKVVVLDLDPWEERSDLEPSIPSEYMLPKKR 2280
 DB 2221 THGKAYDVDMVANLFWGSDVTRIESGKVVVLDLDPWEERSDLEPSIPSEYMLPKKR 2280
 QY 2281 PPPALPAMARPDYNPPLVESWKRDPYQPAITVAGCALPPRRKTPTPPPRRRRRTVGLSEDSI 2340
 DB 2281 PPPALPAMARPDYNPPLVESWKRDPYQPAITVAGCALPPRRKTPTPPPRRRRRTVGLSEDSI 2340
 QY 2341 GVALOOLATKSFQPPSGSDGLSTGAGADSGSQPPPELALSESGISSMPLRGEIG 2400
 DB 2341 GVALOOLATKSFQPPSGSDGLSTGAGADSGSQPPPELALSESGISSMPLRGEIG 2400
 QY 2401 DDDLEBEQVEPQPPPGGVAAPGSDSGSMTCSSEEDSDVCCSMYSMTGALITPCSPBR 2460
 DB 2401 DDDLEBEQVEPQPPPGGVAAPGSDSGSMTCSSEEDSDVCCSMYSMTGALITPCSPBR 2460
 QY 2461 EKLPIPLNSLIRYHNKYYCTTTKSASIRAKKVTEDRMQVLDSDYSDVLKDIKLAASKY 2520
 DB 2461 EKLPIPLNSLIRYHNKYYCTTTKSASIRAKKVTEDRMQVLDSDYSDVLKDIKLAASKY 2520
 QY 2521 TARLLTMEBAQOLTPHSAKSKYFGAKKVRSLSGAVNHHSVMKDLLEDSTPIPTT 2580
 DB 2521 TARLLTMEBAQOLTPHSAKSKYFGAKKVRSLSGAVNHHSVMKDLLEDSTPIPTT 2580
 QY 2581 MAKNEVFCVDPTRKGGKKAARLIYVPLGVRCCKMALYDITOKLPOAVMGASGYFQYSPA 2640
 DB 2581 MAKNEVFCVDPTRKGGKKAARLIYVPLGVRCCKMALYDITOKLPOAVMGASGYFQYSPA 2640
 QY 2641 QREVEFLKMAEKDPMGFSYDTRCFDSTVTERDITREESIYRACSLPEBAHTAHSLTE 2700
 DB 2641 QREVEFLKMAEKDPMGFSYDTRCFDSTVTERDITREESIYRACSLPEBAHTAHSLTE 2700
 QY 2701 RLYVGGPMNSNGQTCGYRCRASGVLTTSMGNTITCYVKALAAACGAAGIAPTMVCGD 2760
 DB 2701 RLYVGGPMNSNGQTCGYRCRASGVLTTSMGNTITCYVKALAAACGAAGIAPTMVCGD 2760
 QY 2761 DLVVISSEQTEDEBNLRAFTTEAMTRYGAPPCDPPREYDELITSCSSNSVALGPOG 2820
 DB 2761 DLVVISSEQTEDEBNLRAFTTEAMTRYGAPPCDPPREYDELITSCSSNSVALGPOG 2820
 QY 2821 RRRYVITRDPTPIARAANETVYHSHSVNSWLGNIIOYAPTIWARMVLMTHPSIIMAQDT 2880
 DB 2821 RRRYVITRDPTPIARAANETVYHSHSVNSWLGNIIOYAPTIWARMVLMTHPSIIMAQDT 2880
 QY 2881 LDONINFEWYGAIVSVSPDLPAIIBRLHGLDAFSLHTYTPHETLRVASALRGLGAPLR 2940
 DB 2881 LDONINFEWYGAIVSVSPDLPAIIBRLHGLDAFSLHTYTPHETLRVASALRGLGAPLR 2940
 QY 2941 AKMSRABARVASLISRGRAAVCGRYLFPMNAVYTKLKLTPLEPARLLDSSMFTVAGGG 3000
 DB 2941 AKMSRABARVASLISRGRAAVCGRYLFPMNAVYTKLKLTPLEPARLLDSSMFTVAGGG 3000
 QY 3001 DIYHSVSRARPRLLLFGLLLLFVGVGLFLLPAR 3033
 DB 3001 DIYHSVSRARPRLLLFGLLLLFVGVGLFLLPAR 3033
 RESULT 2
 AAB31168
 ID AAB31168 standard; protein; 3033 AA.
 XX AAB31168;
 XX
 DT 02-APR-2001 (first entry)
 XX
 DE Amino acid sequence of a hepatitis C virus (HCV) clone genotype 2a.
 XX
 KM Chimeric virus; bovine viral diarrhoea virus; BVDV; hepatitis C virus;
 KM HCV; vaccine; viral inhibitor; antiviral.
 OS Hepatitis C virus.
 XX
 PN WO200075352-A2.

XX 14-DEC-2000.
 PD
 XX
 PF 02-JUN-2000; 2000WO-US015527.
 XX
 XX 04-JUN-1999; 99US-0137817P.
 PR
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Nam J, Buh J, Emerson SU, Purcell RH;
 XX
 DR WPI; 2001-071081/08.
 DR
 DR N-PSDB; AAC86937.
 XX
 PT New nucleic acid comprising a chimeric bovine viral diarrhoea virus genome
 PT in which the (non-)structural region has been replaced by hepatitis C
 PT virus (HCV) genome useful for treating or preventing HCV signs and
 PT symptoms.
 PS
 PS Disclosure; Page 85-97; 97pp; English.
 XX
 CC The specification describes a nucleic acid comprising a chimeric virus
 CC genome, specifically bovine viral diarrhoea virus (BVDV) genome in which
 CC the (non-)structural region has been replaced by the (non-)structural
 CC region of a hepatitis C virus (HCV) genome. The nucleic acids comprising
 CC the chimeric virus and the chimeric virus are useful for identifying cell
 CC lines capable of supporting the replication of these chimeric viruses, in
 CC screening for neutralizing antibodies to HCV of different genotypes, in
 CC the production of HCV-BVDV virions, for the development of inactivated or
 CC attenuated vaccines to prevent HCV-BVDV in a mammal, in studying the
 CC molecular properties of HCV indirectly in vitro, and in identifying
 CC inhibitors of viral enzyme activity which would be useful as antiviral
 CC agents. Formulations or compositions comprising the chimeric viruses may
 CC be used to treat or prevent the signs and symptoms of HCV. The present
 CC sequence is encoded by a HCV clone, which is used to construct chimeric
 CC nucleic acids of the invention
 XX
 SQ Sequence 3033 AA;
 Query Match 100.0%; Score 16120; DB 4; Length 3033;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3033; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSTNPKPORKTKRNTNRPRQDVKFPGGGQIVGGVYLLPFRGRLGVRATRKTSERSQPRG 60
 DB 1 MSTNPKPORKTKRNTNRPRQDVKFPGGGQIVGGVYLLPFRGRLGVRATRKTSERSQPRG 60
 QY 61 RRQPIPKDRRSSTGKSMGKRPYWPPLYGNEGGMAGMLSPRGSRRSWMGPNDRHRSRVNG 120
 DB 61 RRQPIPKDRRSSTGKSMGKRPYWPPLYGNEGGMAGMLSPRGSRRSWMGPNDRHRSRVNG 120
 QY 121 KYIDTLTGGFADLMGYIPVVGAPLGGVVARALAHGVRLVEDGYNFATGNIPEGCSFSIFLLA 180
 DB 121 KYIDTLTGGFADLMGYIPVVGAPLGGVVARALAHGVRLVEDGYNFATGNIPEGCSFSIFLLA 180
 QY 121 KVIDTLTGGFADLMGYIPVVGAPLGGVVARALAHGVRLVEDGYNFATGNIPEGCSFSIFLLA 180
 DB 121 KVIDTLTGGFADLMGYIPVVGAPLGGVVARALAHGVRLVEDGYNFATGNIPEGCSFSIFLLA 180
 QY 181 ILSCITTPVSAEVENISGTGVVNTDCTNDSITWQLQAAVLHVPGCVPECKGNASQCMY 240
 DB 181 ILSCITTPVSAEVENISGTGVVNTDCTNDSITWQLQAAVLHVPGCVPECKGNASQCMY 240
 QY 241 PVSNNVAVQRPALTOGLRTHIDMYVMSATLCSALYVGDLCGVMLAQMFIIVSPQHMF 300
 DB 241 PVSNNVAVQRPALTOGLRTHIDMYVMSATLCSALYVGDLCGVMLAQMFIIVSPQHMF 300
 QY 301 VDDCNCSITPGTITGHRAMAMMMNMSPTATMILAYANRVEVYIIDISGAHWGVFGLA 360
 DB 301 VDDCNCSITPGTITGHRAMAMMMNMSPTATMILAYANRVEVYIIDISGAHWGVFGLA 360
 QY 361 YFSMOGAMAKVVVILLAAAGVDARTHTVGGSAAOQTGRLTSFPDMGPROKIDLVNTNGSM 420
 DB 361 YFSMOGAMAKVVVILLAAAGVDARTHTVGGSAAOQTGRLTSFPDMGPROKIDLVNTNGSM 420
 QY 421 HINRTALNCNDSLATGFIASLFTYTHSFNSGCPERRMSACRSIEARVNGALQYEDNVTN 480
 DB 421 HINRTALNCNDSLATGFIASLFTYTHSFNSGCPERRMSACRSIEARVNGALQYEDNVTN 480

Db 421 HINRTALNCDSLHTGTASLFTYTHSFNSSGCEBERMSACRSIBAFRVGMLQYEDNVTN 480
Qy 481 PEDMRPYCMHYPPROCGVSAKTYCGPVYCFTEPSPVVGTDRLGAPTYTWMGENETDVL 540
Db 481 PEDMRPYCMHYPPROCGVSAKTYCGPVYCFTEPSPVVGTDRLGAPTYTWMGENETDVL 540
Qy 541 LNSTRPPLGSMFGCTMWNSSGYTKTCGAPPCRTRADFNASTDLCPDCCRKPDDTYLK 600
Db 541 LNSTRPPLGSMFGCTMWNSSGYTKTCGAPPCRTRADFNASTDLCPDCCRKPDDTYLK 600
Qy 601 CGSGPMTLPBCLIDYPRLMHYPTVYTI FKIRMYGVGEHRLTAACNTRGDRCNLE 660
Db 601 CGSGPMTLPBCLIDYPRLMHYPTVYTI FKIRMYGVGEHRLTAACNTRGDRCNLE 660
Qy 661 RDRSOLSPLLHSTTEMAILPCSYSDLPALSTGLHLHQNIVDQFMYGSPALTKYIVR 720
Db 661 RDRSOLSPLLHSTTEMAILPCSYSDLPALSTGLHLHQNIVDQFMYGSPALTKYIVR 720
Qy 721 EWTYLLFLLADARVCACLMMLILGQABALEKVLILHAASAACNGFLYFYIIFYAAM 780
Db 721 EWTYLLFLLADARVCACLMMLILGQABALEKVLILHAASAACNGFLYFYIIFYAAM 780
Qy 781 YIKGRVVPATYSLTGMSFSLILALPOQAYAYDASVHGQIGAILVMTLFTLPGYK 840
Db 781 YIKGRVVPATYSLTGMSFSLILALPOQAYAYDASVHGQIGAILVMTLFTLPGYK 840
Qy 841 TLLSRFLMWLCYLLTLGEBANVOEWAPMOYRGGRDGIIMAVAFYPGVVEDITKMLAVL 900
Db 841 TLLSRFLMWLCYLLTLGEBANVOEWAPMOYRGGRDGIIMAVAFYPGVVEDITKMLAVL 900
Qy 901 GPRYLKGALTRVPYFVRAHALRMCTMAHLAGRIVOMALLALGMHTGTYIYDHLTPM 960
Db 901 GPRYLKGALTRVPYFVRAHALRMCTMAHLAGRIVOMALLALGMHTGTYIYDHLTPM 960
Qy 961 SDMAASGLDLAVNEPIIFSPMEKVIYWGASTAACGDIHLGIPVASARLGREVLGPAD 1020
Db 961 SDMAASGLDLAVNEPIIFSPMEKVIYWGASTAACGDIHLGIPVASARLGREVLGPAD 1020
Qy 1021 GYTSKGNLSLAPITAYAQOTRGLLGTIVSMGTGRDKTEQAGEIQVLSTVTSQFLGTSISG 1080
Db 1021 GYTSKGNLSLAPITAYAQOTRGLLGTIVSMGTGRDKTEQAGEIQVLSTVTSQFLGTSISG 1080
Qy 1081 VLMVTVYGAQNKTLAGSRPTQWYMSABEDLVGMBSPPGTKSLBPTCCAVDLIYVTRN 1140
Db 1081 VLMVTVYGAQNKTLAGSRPTQWYMSABEDLVGMBSPPGTKSLBPTCCAVDLIYVTRN 1140
Qy 1141 ADVIPARRRQDKRGALSPRPLSTLKSSSGGPVLCPRGHAVGVPRAVAVCSRGVAKSIDFI 1200
Db 1141 ADVIPARRRQDKRGALSPRPLSTLKSSSGGPVLCPRGHAVGVPRAVAVCSRGVAKSIDFI 1200
Qy 1201 PVETLIDIVTRSPTFSNSTRPAVPQTYQVGYLAHPTSGSKSTKVPVAYAAQYKVLVLP 1260
Db 1201 PVETLIDIVTRSPTFSNSTRPAVPQTYQVGYLAHPTSGSKSTKVPVAYAAQYKVLVLP 1260
Qy 1261 SVAAATLFGAYLSAAGHINPIRGTVTTTGAPITTYTGKPLADOGCAGAYDIIICD 1320
Db 1261 SVAAATLFGAYLSAAGHINPIRGTVTTTGAPITTYTGKPLADOGCAGAYDIIICD 1320
Qy 1321 ECHAVDSSTLIGITVDQAEATAGVRLTVLATATPSPSTYTPHNIIEVALGQEGEIPFY 1380
Db 1321 ECHAVDSSTLIGITVDQAEATAGVRLTVLATATPSPSTYTPHNIIEVALGQEGEIPFY 1380
Qy 1381 GRAIPLSYIKGRHLIFCHSKKKCDLAAALRGMLNSVAYYRGLDVSVIPTQGDVVVA 1440
Db 1381 GRAIPLSYIKGRHLIFCHSKKKCDLAAALRGMLNSVAYYRGLDVSVIPTQGDVVVA 1440
Qy 1441 TDALMTGYTGDPSDIVICNVAVVQVNDPSLDPTTITTOIVPQDAVARSQRGRGTGGR 1500
Db 1441 TDALMTGYTGDPSDIVICNVAVVQVNDPSLDPTTITTOIVPQDAVARSQRGRGTGGR 1500
Qy 1501 GIYRYVSTGERASGMPDSVULCECYDAGAAWYELTTPSETTVRLRAYENTGGLPYCODHLE 1560
Db 1501 GIYRYVSTGERASGMPDSVULCECYDAGAAWYELTTPSETTVRLRAYENTGGLPYCODHLE 1560

Qy 1561 FWEAVFTGLTHIDAFHLSQTKOSGENFAVLTAYQATVCARAKAPPPSNDVMNKCULTRKP 1620
Db 1561 FWEAVFTGLTHIDAFHLSQTKOSGENFAVLTAYQATVCARAKAPPPSNDVMNKCULTRKP 1620
Qy 1621 TLVGPPLLYRLGISTNEVTLTHPTTKYIATCOMQDLVMTSTWLAGVLAAYACIA 1680
Db 1621 TLVGPPLLYRLGISTNEVTLTHPTTKYIATCOMQDLVMTSTWLAGVLAAYACIA 1680
Qy 1681 TGCVCIGRLHINQAVVAPDKEVLYEAFDEMEECASRAALIEEGORIEMLKSKIQGL 1740
Db 1681 TGCVCIGRLHINQAVVAPDKEVLYEAFDEMEECASRAALIEEGORIEMLKSKIQGL 1740
Qy 1741 QOASKOADIOPTVOASMPKYEQFWAKIMMNEISGIOYLAGLSTLPGNPVAVSMAFSA 1800
Db 1741 QOASKOADIOPTVOASMPKYEQFWAKIMMNEISGIOYLAGLSTLPGNPVAVSMAFSA 1800
Qy 1801 LTPSLSTSTTLNLNLGMLASQIAPRAGATGFVYSGVGAAVGSIQGLKVLVDILAGY 1860
Db 1801 LTPSLSTSTTLNLNLGMLASQIAPRAGATGFVYSGVGAAVGSIQGLKVLVDILAGY 1860
Qy 1861 AGISGALVAFKIMSGEKPSEMEDVNLPGIISPGALVYGVICAAILRRHVGBEGAQVM 1920
Db 1861 AGISGALVAFKIMSGEKPSEMEDVNLPGIISPGALVYGVICAAILRRHVGBEGAQVM 1920
Qy 1921 NRLIAFASRGNHVAPTHYTESDASQRYTOLLGSLTTSBLRLRLHMTTEDCPICGGSW 1980
Db 1921 NRLIAFASRGNHVAPTHYTESDASQRYTOLLGSLTTSBLRLRLHMTTEDCPICGGSW 1980
Qy 1981 LRDVMDWCTILTIDRKMTLSLFRKMPLPVSOXOKYGVWAGTGIMTRPCGANIS 2040
Db 1981 LRDVMDWCTILTIDRKMTLSLFRKMPLPVSOXOKYGVWAGTGIMTRPCGANIS 2040
Qy 2041 GNVRLGSMRITGPKCMNIMQSTFPINCYTEBQCVKPAKPNKVAIMRVAASEVAETQH 2100
Db 2041 GNVRLGSMRITGPKCMNIMQSTFPINCYTEBQCVKPAKPNKVAIMRVAASEVAETQH 2100
Qy 2101 GSYYHYITGLTTDNLKVPQCLPSPEFSSWVDQVQIHRPAFTPKPFRFDEVSPCVGLNSPVY 2160
Db 2101 GSYYHYITGLTTDNLKVPQCLPSPEFSSWVDQVQIHRPAFTPKPFRFDEVSPCVGLNSPVY 2160
Qy 2161 GSOLPCDDEPDDVMSMLTDPESHITAEFAARLARSGSPSSASSASQSLAPSLRATCT 2220
Db 2161 GSOLPCDDEPDDVMSMLTDPESHITAEFAARLARSGSPSSASSASQSLAPSLRATCT 2220
Qy 2221 THGKAVDVMDVANLPMGSDVTRIESGSKVVLJDSLDPVVEERSDLEPSIPEBYMLPKKR 2280
Db 2221 THGKAVDVMDVANLPMGSDVTRIESGSKVVLJDSLDPVVEERSDLEPSIPEBYMLPKKR 2280
Qy 2281 PPPALPAMARPDYNPLVESWKRPDYOPATVAGCALPPPRKTPTPPRRRRRTVGLSEDSI 2340
Db 2281 PPPALPAMARPDYNPLVESWKRPDYOPATVAGCALPPPRKTPTPPRRRRRTVGLSEDSI 2340
Qy 2341 GDAOLOALIKSFGCPSPSGDSLSTGAGAADSGSOTPPBELALSTSGSISSMPLLEGEL 2400
Db 2341 GDAOLOALIKSFGCPSPSGDSLSTGAGAADSGSOTPPBELALSTSGSISSMPLLEGEL 2400
Qy 2401 DDLPEBQVEPQPPPGGVAAPGSDSGSWSTSEBEDSVVCCSWSYWTGALITPCSPBE 2460
Db 2401 DDLPEBQVEPQPPPGGVAAPGSDSGSWSTSEBEDSVVCCSWSYWTGALITPCSPBE 2460
Qy 2461 EKLPINPLSNLSLRHNKYYCTTYSASGLRAKCVTFDRMOVLDSYDYLKQIKLAASRY 2520
Db 2461 EKLPINPLSNLSLRHNKYYCTTYSASGLRAKCVTFDRMOVLDSYDYLKQIKLAASRY 2520
Qy 2521 TABLLTMEBACOLTPHSARSXYGGAKEVRLSGRAVNHISKVMWDLLEDESTPPTTI 2580
Db 2521 TABLLTMEBACOLTPHSARSXYGGAKEVRLSGRAVNHISKVMWDLLEDESTPPTTI 2580
Qy 2581 MAKNEVFCVDPTRKGGKARLIVYDPLGYRVEKMAIYDITQKLPQAVMGASGYFOYSRA 2640
Db 2581 MAKNEVFCVDPTRKGGKARLIVYDPLGYRVEKMAIYDITQKLPQAVMGASGYFOYSRA 2640

QY 2641 ORVEFLKMAEKDPMGSDYTRCFDSVTERDITRETSIVRACSLPEEAHTAHSLTE 2700
 Db 2641 QREFFLLKMAEKDPMGSDYTRCFDSVTERDITRETSIVRACSLPEEAHTAHSLTE 2700
 QY 2701 RLTVGGPMFNSKQCTGRCRCSAGVLTSMGNTITCYVALAACGAAGIAPMTLVCGD 2760
 Db 2701 RLTVGGPMFNSKQCTGRCRCSAGVLTSMGNTITCYVALAACGAAGIAPMTLVCGD 2760
 QY 2761 DLVVISGSGTEDEENLRAFTEAMTRYSAPCGDPRPEVDLITSCSNVVALGPGG 2820
 Db 2761 DLVVISGSGTEDEENLRAFTEAMTRYSAPCGDPRPEVDLITSCSNVVALGPGG 2820
 QY 2821 RRRYVITRDPPTTPIAABAMTVHSPVNSMLGNIIOYAPTIMARMVLMTHFSSILMAODT 2880
 Db 2821 RRRYVITRDPPTTPIAABAMTVHSPVNSMLGNIIOYAPTIMARMVLMTHFSSILMAODT 2880
 QY 2881 LDQNLNFMVGAIVSYSPDLPAIERLHGLDAFSLHTYTPHELTREVASALRKLGAAPLR 2940
 Db 2881 LDQNLNFMVGAIVSYSPDLPAIERLHGLDAFSLHTYTPHELTREVASALRKLGAAPLR 2940
 QY 2941 AMKSRARAVRASLSRSGRAAVCGRYLFWMAVTKLKLTPLPBARLLDSSWFTVAGGG 3000
 Db 2941 AMKSRARAVRASLSRSGRAAVCGRYLFWMAVTKLKLTPLPBARLLDSSWFTVAGGG 3000
 QY 3001 DIHVSRRARPRLLFGLLLFVGVGLFLLPAR 3033
 Db 3001 DIHVSRRARPRLLFGLLLFVGVGLFLLPAR 3033

RESULT 3

AA30729
 ID AAB30729 standard; protein; 3033 AA.

AC AAB30729;

DT 02-APR-2001 (first entry)

DE Amino acid sequence of infectious Hepatitis C virus strain HC-J6CH.

KW HCV, HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine.

OS Hepatitis C virus.

PN MO200075338-A2.

PD 14-DEC-2000.

PF 02-JUN-2000; 2000MO-US015446.

PR 04-JUN-1999; 99US-0137693P.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Yanagi M, Bukh J, Emerson SU, Purcell RH;

DR MPI; 2001-061728/07.

DR N-PSDB; AAC86644.

PT Nucleic acid molecule encoding human hepatitis C virus of genotype 2a for
 developing vaccines, for diagnosis of hepatitis C virus and in screening
 assays for identification of antiviral agents.

PS Disclosure; Page 88-99; 167pp; English.

XX The present sequence represents an amino acid sequence of infectious
 CC Hepatitis C virus (HCV) strain HC-J6CH genotype 2a. The HCV
 CC polynucleotide sequence is capable of expressing the virus when
 CC transfected into cells. The HCV protein is useful for assaying candidate
 CC antiviral agents for activity against HCV. Antibodies specific for HCV
 CC polypeptide are useful in prevention and treatment of diseases caused by
 CC HCV in animals, in particular humans. The HCV polypeptides serve as
 CC immunogens in the development of vaccines for preventing HCV in mammals
 CC or as antigens in diagnostic assays for detecting the presence of HCV in

CC biological samples. The HCV polynucleotide is also useful for identifying
 CC cell lines capable of supporting the replication of HCV in vitro and to
 CC produce attenuated viral strains via passage in vitro or in vivo

XX Sequence 3033 AA;

Query Match 100.0%; Score 16120; DB 4; Length 3033;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3033; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKRQKTKRNTNRPRQDYKPGGQIVGGVLLPRRPRLGVRATRKTSERSQPRG 60
 Db 1 MSTNPKRQKTKRNTNRPRQDYKPGGQIVGGVLLPRRPRLGVRATRKTSERSQPRG 60
 QY 61 RROPIDKDRRSYSGSKSGKGYPMPLYGNGGLGAGPLSPRGRBPMGPDPHRRRNWG 120
 Db 61 RROPIDKDRRSYSGSKSGKGYPMPLYGNGGLGAGPLSPRGRBPMGPDPHRRRNWG 120
 QY 121 KYIDTLCGFADLMGYIPVVGAPLGGVARALAHGVLEEDGVNFAFGNPGCSFIPLLA 180
 Db 121 KYIDTLCGFADLMGYIPVVGAPLGGVARALAHGVLEEDGVNFAFGNPGCSFIPLLA 180
 QY 181 LLSCTTPVSAAEVKNISITGYVNTDCTNDSITWQLOAAVLHVPCVPCERKYNASQCM 240
 Db 181 LLSCTTPVSAAEVKNISITGYVNTDCTNDSITWQLOAAVLHVPCVPCERKYNASQCM 240
 QY 241 PVSBNVAVORPQALTOGLRTHIDMVVMSATLCSALYVGLCGGVMLAQMFIYSPQHMF 300
 Db 241 PVSBNVAVORPQALTOGLRTHIDMVVMSATLCSALYVGLCGGVMLAQMFIYSPQHMF 300
 QY 301 VODCNCSITPGTITGHRMAMDMNMNMPATMILAYAMPREVIDIISGAHWVFGFLA 360
 Db 301 VODCNCSITPGTITGHRMAMDMNMNMPATMILAYAMPREVIDIISGAHWVFGFLA 360
 QY 361 YFSMOGAMAKVVVILLAAVGDARTHTVGSAAQTGRLTSLPDMGPROKIQLVNTNSW 420
 Db 361 YFSMOGAMAKVVVILLAAVGDARTHTVGSAAQTGRLTSLPDMGPROKIQLVNTNSW 420
 QY 421 HINRTALNCNDSLHNGFIASLFTYHSFNSSGCEPERSACRSIEAFVWGALQYEDNVN 480
 Db 421 HINRTALNCNDSLHNGFIASLFTYHSFNSSGCEPERSACRSIEAFVWGALQYEDNVN 480
 QY 481 PEDMRPYCWHYPPROCGVVSAAKTVCGPVYCFTPSPVVVGTTRTGLAPYTWGNETDVL 540
 Db 481 PEDMRPYCWHYPPROCGVVSAAKTVCGPVYCFTPSPVVVGTTRTGLAPYTWGNETDVL 540
 QY 541 LNSTRPPLGSPGCTWNNSSGYTKTCGAPPCRTADFNASTDLCEPTDCEFRKHPDTTYLK 600
 Db 541 LNSTRPPLGSPGCTWNNSSGYTKTCGAPPCRTADFNASTDLCEPTDCEFRKHPDTTYLK 600
 QY 601 CGSGPWLTPRCILIDPYRLMHIPTVYNTYIFKIRMYGVGVERRLTAAACNFTGDRCNLED 660
 Db 601 CGSGPWLTPRCILIDPYRLMHIPTVYNTYIFKIRMYGVGVERRLTAAACNFTGDRCNLED 660
 QY 661 RDRSOLSPLLHSTTEMAILPCSYSDLPALSTGLHLHONIIVVOVFYSGPALTKYIVM 720
 Db 661 RDRSOLSPLLHSTTEMAILPCSYSDLPALSTGLHLHONIIVVOVFYSGPALTKYIVM 720
 QY 721 EWWIILFLLLADARVACIMMLILGQAEALAEKLVILHAASASNGEFLYVIFVFAAM 780
 Db 721 EWWIILFLLLADARVACIMMLILGQAEALAEKLVILHAASASNGEFLYVIFVFAAM 780
 QY 781 YIKGRVPLATYSLTGLMSFSLILLALPOQAAAYASVHGOIGALLVMTLFTLTPGXK 840
 Db 781 YIKGRVPLATYSLTGLMSFSLILLALPOQAAAYASVHGOIGALLVMTLFTLTPGXK 840
 QY 841 TLISFPLMWLCYLLTGEMVQEMAPPMQVRGGRDGIWAVALFYPGVVFDTIKMLAVAL 900
 Db 841 TLISFPLMWLCYLLTGEMVQEMAPPMQVRGGRDGIWAVALFYPGVVFDTIKMLAVAL 900
 QY 901 GPAYLLKGLATVPYFVRHALLRMCWTARHLAGRYVOMALLALGRWTGYIYDHLTPM 960
 Db 901 GPAYLLKGLATVPYFVRHALLRMCWTARHLAGRYVOMALLALGRWTGYIYDHLTPM 960

Qy 961 SDMAAGLRLAVAVEPILFSPMEKVIWGAETAAAGDILHGLPVSAARLGRREVLLGPAD 1020
Db 961 SDMAAGLRLAVAVEPILFSPMEKVIWGAETAAAGDILHGLPVSAARLGRREVLLGPAD 1020
Qy 1021 GYTSKGSLLAPITAYAOQTRGLGITVSMGTGDKTEQAGEIQVLTSTVQSPFLGISISG 1080
Db 1021 GYTSKGSLLAPITAYAOQTRGLGITVSMGTGDKTEQAGEIQVLTSTVQSPFLGISISG 1080
Qy 1081 VLMVTYVAGNKTLAGSRGPVOMYSABEDLVGMPSPGPKTSLEPTCGADVLYVTRN 1140
Db 1081 VLMVTYVAGNKTLAGSRGPVOMYSABEDLVGMPSPGPKTSLEPTCGADVLYVTRN 1140
Qy 1141 ADVIPARRRDKGALLSPRLSLTKSGSGGPVLCPRGHAUVGFRAAVCSRGVAKSIDP1 1200
Db 1141 ADVIPARRRDKGALLSPRLSLTKSGSGGPVLCPRGHAUVGFRAAVCSRGVAKSIDP1 1200
Qy 1201 PVETLIDIVTRSPFSDNSTPPAVQTVQVYLHAPTSKSTKVPVAAOQYKVLVLP 1260
Db 1201 PVETLIDIVTRSPFSDNSTPPAVQTVQVYLHAPTSKSTKVPVAAOQYKVLVLP 1260
Qy 1261 SVAAITLFGAYLSKAHGIMPVIRGVRVTTGAPITVSTYGFKAODGCGAGAVDITICD 1320
Db 1261 SVAAITLFGAYLSKAHGIMPVIRGVRVTTGAPITVSTYGFKAODGCGAGAVDITICD 1320
Qy 1321 ECHAVDSTTILIGITVLDQAEAGVRLTVLATATPPGSVTPPHNIEVALQGEGETPFY 1380
Db 1321 ECHAVDSTTILIGITVLDQAEAGVRLTVLATATPPGSVTPPHNIEVALQGEGETPFY 1380
Qy 1381 GRALPSTYIKGGRHLIFCHSKKCCDELAALRGMLNSVAYRGLDVSVIPGQDVVVA 1440
Db 1381 GRALPSTYIKGGRHLIFCHSKKCCDELAALRGMLNSVAYRGLDVSVIPGQDVVVA 1440
Qy 1441 TDALMTGYTGDPSVIDICNVAVTOVDFSLDPFTITTOIVPODAVRSORRGTRGRL 1500
Db 1441 TDALMTGYTGDPSVIDICNVAVTOVDFSLDPFTITTOIVPODAVRSORRGTRGRL 1500
Qy 1501 GIYRYVSTGERASGMPDSVVLCECYDAGAAAYELTPSETTVRLRAYENTPGLPYCODHLE 1560
Db 1501 GIYRYVSTGERASGMPDSVVLCECYDAGAAAYELTPSETTVRLRAYENTPGLPYCODHLE 1560
Qy 1561 FMEVAFGLTHIDAHFISQTKSGENPAVLYTAYATYCARAKAPRPBMDVMWKLTRLP 1620
Db 1561 FMEVAFGLTHIDAHFISQTKSGENPAVLYTAYATYCARAKAPRPBMDVMWKLTRLP 1620
Qy 1621 TLVGPETLYRISGVTEVTLTHPTKYIATCMQADLEVMTSTVNLAGVLAAYACLA 1680
Db 1621 TLVGPETLYRISGVTEVTLTHPTKYIATCMQADLEVMTSTVNLAGVLAAYACLA 1680
Qy 1681 TGCVCIIIGRLHINQRAVVAPOKEVLYEAFDEMEBCASRAALIBEGQRIAMELSKIOGLL 1740
Db 1681 TGCVCIIIGRLHINQRAVVAPOKEVLYEAFDEMEBCASRAALIBEGQRIAMELSKIOGLL 1740
Qy 1741 QOASQKODIOPVQOASMPKVEQFMAGMMNFISGIOTYLAGLSTLPENPAVAMAFSA 1800
Db 1741 QOASQKODIOPVQOASMPKVEQFMAGMMNFISGIOTYLAGLSTLPENPAVAMAFSA 1800
Qy 1801 LTPSLSTSTTILNIIIGMLASOIAPPAGATGFVSGVGAAYVSGIGAGVLYVDIILAGYG 1860
Db 1801 LTPSLSTSTTILNIIIGMLASOIAPPAGATGFVSGVGAAYVSGIGAGVLYVDIILAGYG 1860
Qy 1861 AGISGALVAFKIMGKESKEMEDVNLPGILSPGALVVGVICAILRRHVGPGGAVQM 1920
Db 1861 AGISGALVAFKIMGKESKEMEDVNLPGILSPGALVVGVICAILRRHVGPGGAVQM 1920
Qy 1921 NRLIAFASRGNHVAPTHVVTESDASQRYTQLLSLITTSILRLAHMTTEDCP1PCGGSW 1980
Db 1921 NRLIAFASRGNHVAPTHVVTESDASQRYTQLLSLITTSILRLAHMTTEDCP1PCGGSW 1980
Qy 1981 LRDVMDVWCTILVDFKQMLTSLCKPKMPGLPFVSCQKGYGVWAGTGIINTRCGCGANIS 2040
Db 1981 LRDVMDVWCTILVDFKQMLTSLCKPKMPGLPFVSCQKGYGVWAGTGIINTRCGCGANIS 2040

Qy 2041 GNVRLGSKRITGPKTCOMINMOGTFPINCYTEGQVCPKAPNFKVAILMRYAASEVAETQH 2100
Db 2041 GNVRLGSKRITGPKTCOMINMOGTFPINCYTEGQVCPKAPNFKVAILMRYAASEVAETQH 2100
Qy 2101 GSYHYITGTLTNNLKVPCQLBEPFESWVDGQVIRFAPTPPFRDREVSFCVGLNSPVY 2160
Db 2101 GSYHYITGTLTNNLKVPCQLBEPFESWVDGQVIRFAPTPPFRDREVSFCVGLNSPVY 2160
Qy 2161 GSQLPCEPDPDVLMSMLTDPSSHITAEYAAARLARGSPSSASSASAGLSAPSRATCT 2220
Db 2161 GSQLPCEPDPDVLMSMLTDPSSHITAEYAAARLARGSPSSASSASAGLSAPSRATCT 2220
Qy 2221 THGKAVDVMVDANLPMGSDVTRIESGKVVVLDLDPWVEERSDLEPSIPSEYMLPKKR 2280
Db 2221 THGKAVDVMVDANLPMGSDVTRIESGKVVVLDLDPWVEERSDLEPSIPSEYMLPKKR 2280
Qy 2281 PPPALPAMARPDYNNPVLVESMKRPDYOPATTVAGCALLPPRKPTPTPPRRRRVGLSEDSI 2340
Db 2281 PPPALPAMARPDYNNPVLVESMKRPDYOPATTVAGCALLPPRKPTPTPPRRRRVGLSEDSI 2340
Qy 2341 GDALQOLAIKSRGQPPSGDSGLSTGAGAADSGSOTPPDELALSETGISSNMPLBGLG 2400
Db 2341 GDALQOLAIKSRGQPPSGDSGLSTGAGAADSGSOTPPDELALSETGISSNMPLBGLG 2400
Qy 2401 DPLBEBQVEPOPPOGVAABGSDSGSWSTCSSEBDSVVCSSMSYMTGALITPCSPBE 2460
Db 2401 DPLBEBQVEPOPPOGVAABGSDSGSWSTCSSEBDSVVCSSMSYMTGALITPCSPBE 2460
Qy 2461 EKLPTNPNSLILRHNKRYCTTTKSASIRAKKTFPDRQVUDSYVDSVLKDIKLAASRY 2520
Db 2461 EKLPTNPNSLILRHNKRYCTTTKSASIRAKKTFPDRQVUDSYVDSVLKDIKLAASRY 2520
Qy 2521 TARLITMEBAQOLTPPHSARSKYGFCAKEVRSLSGRAVNHISWKMDLLEDBETP1PTTI 2580
Db 2521 TARLITMEBAQOLTPPHSARSKYGFCAKEVRSLSGRAVNHISWKMDLLEDBETP1PTTI 2580
Qy 2581 MAKNEVFCVDPKGGKKAARLIVYDPLGVRCEKALYDITOKLPOAVMGASYGFOYSPA 2640
Db 2581 MAKNEVFCVDPKGGKKAARLIVYDPLGVRCEKALYDITOKLPOAVMGASYGFOYSPA 2640
Qy 2641 QREVEFLKMAKXKOPMGFSYDTRCPDSVTERDRI RTEBSIYRAGSLPEBAHTALHSLTE 2700
Db 2641 QREVEFLKMAKXKOPMGFSYDTRCPDSVTERDRI RTEBSIYRAGSLPEBAHTALHSLTE 2700
Qy 2701 RLYVGGPMFNSKQOTCYRCRASGLTSMGNITTCYKALAAKAAAGIIAPTMLVCGD 2760
Db 2701 RLYVGGPMFNSKQOTCYRCRASGLTSMGNITTCYKALAAKAAAGIIAPTMLVCGD 2760
Qy 2761 DLVVISSEQGTEDERNILRAFTBAMTRYSAAPGDPREPRYDLELITSCSNVSVALGPOG 2820
Db 2761 DLVVISSEQGTEDERNILRAFTBAMTRYSAAPGDPREPRYDLELITSCSNVSVALGPOG 2820
Qy 2821 RRRYLTTRPPTPIARAAMEYTRHSPVNSWLNIIIOYARTIARAVLMTHPFSIILAAOT 2880
Db 2821 RRRYLTTRPPTPIARAAMEYTRHSPVNSWLNIIIOYARTIARAVLMTHPFSIILAAOT 2880
Qy 2881 LDONLNFEMYGAVYSPLDPAIIERLHGLDAFSLHTYTPHELTRVSAALKLQAPPLR 2940
Db 2881 LDONLNFEMYGAVYSPLDPAIIERLHGLDAFSLHTYTPHELTRVSAALKLQAPPLR 2940
Qy 2941 AMKSRARAVRASILISRGRAAVCGRYLFNMAVYTKLKLTLPDPEARLLDLSWFTVAGCGG 3000
Db 2941 AMKSRARAVRASILISRGRAAVCGRYLFNMAVYTKLKLTLPDPEARLLDLSWFTVAGCGG 3000
Qy 3001 DIYHSVSRARPRLLIFGLLLFVGCLFLLPPAR 3033
Db 3001 DIYHSVSRARPRLLIFGLLLFVGCLFLLPPAR 3033

RESULT 4
AAK3538
ID AAK3538 standard; proteoin; 3033 AA.
XX

AC	AAR33538;
XX	
DT	27-AUG-2003 (revised)
DT	25-MAR-2003 (revised)
DT	01-JUL-1993 (first entry)
XX	
DE	NANBH virus strain HC-J6 protein.
XX	
KM	Non A non B hepatitis virus; plasma.
OS	Non-A.
XX	
OS	non-B hepatitis virus.
XX	
PN	EP532167-A2.
XX	
PD	17-MAR-1993.
XX	
PF	30-JUL-1992; 92EP-00306952.
XX	
PR	09-AUG-1991; 91JP-00287402.
PR	05-DEC-1991; 91JP-00360441.
XX	
PA	(IMMO) IMMUNO JAPAN INC.
XX	
PI	Okamoto H, Nakamura T;
XX	
DR	WPI, 1993-087166/11.
DR	P-PsDB; AAQ38218.
XX	
PT	polynucleotide(s), polypeptide(s) and antibodies of NANBH virus - useful for detecting NANBH, as a vaccine and for screening blood samples.
XX	
PS	Claim 5; Page 38-52; 93pp; English.
XX	
CC	RNA was isolated from the plasma of human patients positive for NANBH virus (strain HC-J6) and was subjected to reverse transcription to produce cDNA. The resulting cDNA was amplified by PCR, and nucleic acid sequences determined by analysis of both clones from the cDNA library and clones obcd. by PCR amplification (36 clones in total). The NANBH HC-J6 genome was found to contain an open reading frame encoding a polypeptide precursor of 3033 amino acid residues. See also AAR33539 and AAR33214.
CC	(Updated on 25-MAR-2003 to correct PN field.) (updated on 27-AUG-2003 to correct OS field.)
CC	
XX	
SQ	Sequence 3033 AA;
Query Match	98.2%; Score 15824; DB 2; Length 3033;
Best Local Similarity	97.8%; Pred. No. 0;
Matches 2966; Conservative 30; Mismatches 37; Indels 0; Gaps 0	
OY	1 MSTNEKPKORTKRNTRRRPODYKFPGGGQIVGGVYLPPRGPRLGVRATRTKSERSQPRG 60
Db	1 MSTNKKPQRKTGRNTNRRRPODYKFPGGGQIVGGVYLPPRGPRLGVRATRTKSERSQPRG 60
OY	61 RRQPIPKORRSTGSKMGKPGFYWPPLYGNELGMAGWLISPRGSRPSPGNDPRHRSRVNG 120
Db	61 RROPIDKRRSSTGSKMGKPGFYWPPLYGNELGMAGWLISPRGSRPSPGNDPRHRSRVNG 120
OY	121 KYIDILTLTGCFADLMGYIIPVVGAPLGGVAALAHGVNLVDGVNFAFGNLPGCSFSIFLLA 180
Db	121 KYIDILTLTGCFADLMGYIIPVVGAPLGGVAALAHGVNLVDGVNFAFGNLPGCSFSIFLLA 180
OY	181 LLSCTITPEPSAAEVENKNISTGVWVTNDCTNDSITWQLQAVALHVPGVCPECKYGNAOCMI 240
Db	181 LLSCTITPEPSAAEVENKNISTGVWVTNDCTNDSITWQLQAVALHVPGVCPECKYNTSRCHI 240
OY	241 PVSFNVAVORPGALTQGLRTHIDVMVMSATLCSALYYGDLCGGVMLAQMFIIVSPQHMF 3000
Db	241 PVSFNVAVOPGALTQGLRTHIDVMVMSATLCSALYYGDLCGGVMLAQMFIIVSPQHMF 3000
OY	301 VODCNCSIYPGTITIGHRMADMNMNSPTRATMTILAAMVPEVIIDIIGAHHGVMFGLA 360
Db	301 VODCNCSIYPGTITIGHRMADMNMNSPTRATMTILAAMVPEVIIDIIGAHHGVMFGLA 360

QY	36L	PFMSOGAAKVVVILLILAAGVAPRTHYTGSSAAQTTGTLTSPJEDMOPROKIOUVNNGSM	420
Db	36L	YFSMGAAKVVVILLILAAGVDAQHTHYTGSSAHHARHTITGMSFGAROKIOUINNGSM	420
QY	421	HINRTALNCNLSLHTGFASLEFYTHSFNMSGCPERNMSACRSIEAPRVGALQYEDNVTN	480
Db	421	HINRTALNCNDSLHTGFASLIFYTHSFNMSGCPERNMSACRSIEAPRVGALQYEDNVTN	480
QY	461	PEDMPCYCMHPPROCGVSVASATVCVGYCFPPSPVUVGTORLGAPTYTWGENEDVEL	540
Db	461	PEDMPCYCMHPPROCGVSVASVCGPYCFPPSPVUVGTORLGAPTYTWGENEDVEL	540
QY	541	INSTRPPLSGMFCGCMWMSGGYTKTCGAPPCRTRADFNASTDLLCTDPCFRKHPDTTYK	600
Db	541	INSTRPPOGSMFCGCMWMSGTGTYTCGAPPCIRADFNASMDLCTDPCFRKHPDTTYK	600
QY	601	CGSGFWLTPRCLIDYPRYLMHYPCIVNVTIFKIRMYVGVVERHLTAACNFTRGDRCNLED	660
Db	601	CGSGFWLTPRCLIDYPRYLMHYPCIVNVTIFKIRMYVGVVERHLTAACNFTRGDRCNLED	660
QY	661	RDRSOLSPLSHSTTEMALIPCSYDLPALSTELHMQIVOVOPMGYSPLTKIYBM	720
Db	661	RDRSOLSPLSHSTTEMALIPCTYSDLPALSTELHMQIVOVOPMGYSPLTKIYBM	720
QY	721	EWVILLFLLADARVACILMMLILGOAEBALEKVIILHAASACNGEYFVIFPVAAM	780
Db	721	EWVILLFLLADARVACIMMLILGOAEBALEKVLILHAASACNGEYFVIFPVAAM	780
QY	781	YIKGVNVPPLATYSLTGLMSFSLILLALPOOANAYASVHGOGALLWNITLFTLTPGK	840
Db	781	YIKGVNVPPLATYSLTGLMSFGLILLALPOOANAYASVHGOGALLVITLFTLTPGK	840
QY	841	TLSRFLMMLCYLLTLTGEAMVOEMAPPMQVNRGGRGIIIAVAIVFYGVVPFDTIKMLAVAL	900
Db	841	TLSRFLMMLCYLLTLTLEAMVOEMAPPMQVNRGGRGIIIAVAIFCGGVVPFDTIKMLAVAL	900
QY	901	GPAYILKQALTRVPYFVRALHLLRMCTMARHLAGRYOMALLAGRMGTGTYIYDHLTPM	960
Db	901	GPAYILKQALTRVPYFVRALHLLRMCTMARHLAAGRYOVMALLAGRMGTGTYIYDHLTPM	960
QY	961	SDMAASGRDLAVAAVEPIIFSPMEKKVIYMGSETAACGDIHJGLPVSARLGBEVLGPPAD	102
Db	961	SDMAANGRDLAVAAVEPIIFSPMEKKVIYMGSETAACGDIHJGLPVSARLGBEVLGPPAD	102
QY	1021	GYSKGSMLAPITAYAOQTRGLGTVIVSMTGDRKTEBAGEIOVLSTVYSGFLGTSISG	108
Db	1021	GYSKGSMLAPITAYAOQTRGLGTVIVSMTGDRKTEBAGEIOVLSTVYSGFLGTTISG	108
QY	1061	VLMVYTHGAKNTTAGSGRPYTOQMSASAEGLVWGPSPGCTSLSECTCGAVIDLYVTBN	114
Db	1061	VLMVYTHGAKNTTAGSGRPYTOQMSASABGDLWGPSPGCTSLSECTCGAVIDLYVTBN	114
QY	1141	ADVIPARRGRKRGALLSPRLSTYKGSAGGVVLCPRGHAVGVFPAAVCSRGVAKSIDET	120
Db	1141	ADVIPARRGRKRGALLSPRLSTLKGSAGGVVLCPRGHAVGVFPAAVCSRGVAKSIDET	120
QY	1201	PVETLIDITRSPTPESDNSTPPAVPOTYOVGYLHAAPTGSKSTKVPVAAOQKVLVLP	126
Db	1201	PVETLIDITRSPTPESDNSTPPAVPOTYOVGYLHAAPTGSKSTKVPVAAOQKVLVLP	126
QY	1261	SVAATLFGGAYLSKSHGINPNIRTVRYTTGCAPIITYSTYSGKFLADGCGAGAYDIIICD	132
Db	1261	SVAATLFGGAYLSKSHGINPNIRTVRYTTGCAPIITYSTYSGFLADGCGAGAYDIIICD	132
QY	1321	ECHAVDSTTIIIGITVDQDAETAGVRLTVLATATPPGSVTTBPHNIEVALCOGEIEIPY	138
Db	1321	ECHAVDSTTIIIGITVDQDAETAGVRLTVLATATPPGSVTTBPHNIEVALCOGEIEIPY	138
QY	1361	GRAIPFLSYIKGRHILPCHSKKKCODELAALRGMLNVAAYVRGIDVSYIPIQSGDVVVVA	144
Db	1361	GRAIPFLSYIKGRHILPCHSKKKCODELAALRGMLNVAAYVRGIDVSYIPIQSGDVVVVA	144

```

QY 1441 TDALMTGYTGDPSIDICNNAVTVVDPSLDPTFTTQIIVPDODASRSQRGRGTGRGL 1500
D 1441 TDALMTGFTGDPDVIDICNNAVTVVDPSLDPTFTTQIIVPDODASRSQRGRGTGRGL 1500
QY 1501 GIYVYVSTGRASGMFDSVYLCECYDAGAAWELTSEETVRLRAYVNTGGLVYCODHLE 1560
D 1501 GIYVYVSTGRASGMFDSVYLCECYDAGAAWELTSEETVRLRAYVNTGGLVYCODHLE 1560
QY 1561 FMEAVFTGLTHIDAHFLSQTQSGENFAVLTAYOATVACARAKAPPSMDVMWKLTRLRK 1620
D 1561 FMEAVFTGLTHIDAHFLSQTQSGENFAVLTAYOATVACARAKAPPSMDVMWKLTRLRK 1620
QY 1621 TLVGPPTLLYRLSGVTNEVTLTHPTKYIATCMQADIEVMTSTVLAAGVLAAYACLA 1680
D 1621 TLVGPPTLLYRLSGVTNEVTLTHPTKYIATCMQADIEVMTSTVLAAGVLAAYACLA 1680
QY 1681 TGCCCTIGRLHINORAVAPDKEVLYEAFDMEBCASRALIERGQRIAEMLSKIOGL 1740
D 1681 TGCCCTIGRLHINORAVAPDKEVLYEAFDMEBCASRALIERGQRIAEMLSKIOGL 1740
QY 1741 QQASKQADIOPTVOASMPKVEQFMAKHMNFISGIOYLAGLSTLPGNPAVASMMAFSA 1800
D 1741 QQASKQADIOPTVOASMPKVEQFMAKHMNFISGIOYLAGLSTLPGNPAVASMMAFSA 1800
QY 1801 LTPSLSTSTTLNLITLGSWLASQIAPAGATGVVSGLVGAANGSICLGYLVYDIIAGYG 1860
D 1801 LTPSLSTSTTLNLITLGSWLASQIAPAGATGVVSGLVGAANGSICLGYLVYDIIAGYG 1860
QY 1861 AGISGALVAFKIMGSRKPSMEDVNLPLGLSPGALVGVYCAAILRRHGPBGAVOM 1920
D 1861 AGISGALVAFKIMGSRKPSMEDVNLPLGLSPGALVGVYCAAILRRHGPBGAVOM 1920
QY 1921 NRLIAFASRGNHVAPTHYVTESDASQRTQLLSLTSLRLHNMITEDECPICGSGW 1980
D 1921 NRLIAFASRGNHVAPTHYVTESDASQRTQLLSLTSLRLHNMITEDECPICGSGW 1980
QY 1981 LRDVMDVCTILDPKMWTLSKLPKMPGLPFVSCQGYGVMAAGTGMTTRPCGANIS 2040
D 1981 LRDVMDVCTILDPKMWTLSKLPKMPGLPFVSCQGYGVMAAGTGMTTRPCGANIS 2040
QY 2041 GNVVLGSMRTITGPKTCNNINOCPTINCYTCGOCVPAPAPFKAIIRVASEVAETOH 2100
D 2041 GNVVLGSMRTITGPKTCNNINOCPTINCYTCGOCVPAPAPFKAIIRVASEVAETOH 2100
QY 2101 GSYHYITGLTTDNLKVFCQLPSPEFPMVDGQIHRAPTPKPFPRDEVFVCVGLNSFV 2160
D 2101 GSYHYITGLTTDNLKVFCQLPSPEFPMVDGQIHRAPTPKPFPRDEVFVCVGLNSFV 2160
QY 2161 GSQLPCDPEPDTVLTSMLTDPHSITAEARRLARSGSPSEASSSASQLSAPSLRATCT 2220
D 2161 GSQLPCDPEPDTVLTSMLTDPHSITAEARRLARSGSPSEASSSASQLSAPSLRATCT 2220
QY 2221 THGKAYVDVMDVDALEPMQGDVTRIESGSKVYVLDLPMVEEREDLPSPSEVYMLPKR 2280
D 2221 THGKAYVDVMDVDALEPMQGDVTRIESGSKVYVLDLPMVEEREDLPSPSEVYMLPKR 2280
QY 2281 FPPALPMARPDYNPPLVESWKRDPYOPATVAGCALPPRKTPTPPRRRTVGLSBSI 2340
D 2281 FPPALPMARPDYNPPLVESWKRDPYOPATVAGCALPPRKTPTPPRRRTVGLSBSI 2340
QY 2341 GDALQQLAIKSGQPPSGDSGLSTGAGADSGSQTPPDLIASSETGISMPRLBGEIG 2400
D 2341 GDALQQLAIKSGQPPSGDSGLSTGAGADSGSQTPPDLIASSETGISMPRLBGEIG 2400
QY 2401 DPLLEPGVEPQPPQGVAAFGSDSGSWSTCSEEDSVVCCSNYSWTGALLTPCGPEE 2460
D 2401 DPLLEPGVEPQPPQGVAAFGSDSGSWSTCSEEDSVVCCSNYSWTGALLTPCGPEE 2460
QY 2461 EKLPIPLSNLSLRYHNKVCTTTSKASIRAKKTPFRMQVLJDYYSVLDKIDLAASKV 2520
D 2461 EKLPIPLSNLSLRYHNKVCTTTSKASIRAKKTPFRMQVLJDYYSVLDKIDLAASKV 2520
QY 2521 TALLTMBEACQLTPPHSARSKYGFGAKEVRSLSGRAVNHISVWKDLLEDSPTPIPTTI 2580

```

```

D 2521 TALLTMBEACQLTPPHSARSKYGFGAKEVRSLSGRAVNHISVWKDLLEDSPTPIPTTI 2580
QY 2581 MAKNEVFCVDPTRKGGKKAARLVYPDLCGVRVEKMAVDIOTKLPOAVMGASGYFOYSPA 2640
D 2581 MAKNEVFCVDPTRKGGKKAARLVYPDLCGVRVEKMAVDIOTKLPOAVMGASGYFOYSPA 2640
QY 2641 QREVFLLKAMARKDPMGFSYDTRCFDSTVTERDTRTESIYRACSLPEBAHTAHSLTE 2700
D 2641 QREVFLLKAMARKDPMGFSYDTRCFDSTVTERDTRTESIYRACSLPEBAHTAHSLTE 2700
QY 2701 RLTVGGPMFNSKGOTGRCRCAAGVLTSMGNTITCYKALAACKAAGIAPMLVCGD 2760
D 2701 RLTVGGPMFNSKGOTGRCRCAAGVLTSMGNTITCYKALAACKAAGIAPMLVCGD 2760
QY 2761 DLVVISSEQGTEDERNLAFBAMTRYSAPGDPREPYDELITSCSNYSVALGPG 2820
D 2761 DLVVISSEQGTEDERNLAFBAMTRYSAPGDPREPYDELITSCSNYSVALGPG 2820
QY 2821 RRRYVLTBDPTPIARAAMETVRHSPVNSWLNIIQVPTIWARVMTHPFSILMAODT 2880
D 2821 RRRYVLTBDPTPIARAAMETVRHSPVNSWLNIIQVPTIWARVMTHPFSILMAODT 2880
QY 2881 LDONLNFEMYGAVYSPLDPAIIRLGLDAFSLHTTTPHETRVASALRKAGAPPR 2940
D 2881 LDONLNFEMYGAVYSPLDPAIIRLGLDAFSLHTTTPHETRVASALRKAGAPPR 2940
QY 2941 AMKSRARAVASLISRGRAACGRYLFMAVAKTLKTLPLPEARLLDLSWFTVAGAGG 3000
D 2941 AMKSRARAVASLISRGRAACGRYLFMAVAKTLKTLPLPEARLLDLSWFTVAGAGG 3000
QY 3001 DIHVSARAPRLLFGLLLFVGVGLFLLPAR 3033
D 3001 DIHVSARAPRLLFGLLLFVGVGLFLLPAR 3033

RESULT 5
ADC42917
ID ADC42917 standard; protein; 2940 AA.
XX
AC ADC42917;
XX
DT 18-DEC-2003 (first entry)
XX
DE Hepatitis C virus protein.
XX
KW viral; YPXU; cellular protein Ap-50; hepatitis C virus; HCV;
KW herpes simplex virus; HSV; UL42 protein; variola virus A10L;
KW vaccinia virus vltion core protein p4a;
KW human parainfluenza virus haemagglutinin-neuraminidase; HSP-1;
KW viral budding; viral infectivity.
XX
OS Hepatitis C virus.
XX
PN MO2003017943-A2.
XX
PD 06-MAR-2003.
XX
PF 22-AUG-2002; 2002W0-US027066.
XX
PR 22-AUG-2001; 2001US-0314182P.
XX
PA (MYRI-) MYRIAD GENETICS INC.
XX
PI Morham S, Zavitz K, Hobden A;
XX
DR MPI; 2003-371696/35.
XX
PT Novel peptide, comprising a contiguous amino acid sequence of a viral
PT protein, capable of binding to a region of cellular protein Ap-50 1a
PT useful for treating viral infections.
XX
XX Example 3; SEQ ID NO 13; 60pp; English.

```

XX The invention relates to a novel isolated peptide comprising a contiguous
CC amino acid sequence of 7-30 amino acid residues of a viral protein, where
CC the contiguous amino acid sequence encompasses the YPKL motif of the
CC viral protein, and the 7-30 amino acid viral protein is capable of
CC binding a region including the amino acid residues 121-435 of cellular
CC protein AP-50. The isolated 7-30 amino acid viral protein is selected
CC from hepatitis C virus (HCV) polypeptide, herpes simplex virus (HSV) UL42
CC protein, varicella virus A10L protein, vaccinia virus virion core protein
CC P4a, and human parainfluenza virus haemagglutinin-neuraminidase, where
CC the contiguous amino acid sequence encompasses the YPKL motif of the viral
CC protein. The peptides of the invention are useful in the manufacture of a
CC medicament for treating viral infections caused by a virus selected from
CC HCV, HSP-1 and varicella virus. The peptides are also useful for inhibiting
CC viral budding from virus-infected cells and/or inhibiting viral
CC infectivity, thus inhibiting viral propagation in the cells. This
CC sequence represents the protein of a Hepatitis C virus used in the
CC treatment of viral infections of the invention.

XX Sequence 2940 AA;

Query Match 97.1%; Score 15645; DB 7; Length 2940;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2940; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKRQKTKXNTRRPPDVKFPGGQGVGGVYLPRGRPLGYRATKTSERSQPRG 60
DB 1 MSTNPKRQKTKXNTRRPPDVKFPGGQGVGGVYLPRGRPLGYRATKTSERSQPRG 60
QY 61 RRQPIKDRRSTGSKGKGPWPLYGNBGLGNAWMLSPRSGRPSGPNDRPHRRSNV 120
DB 61 RRQPIKDRRSTGSKGKGPWPLYGNBGLGNAWMLSPRSGRPSGPNDRPHRRSNV 120
QY 121 KYIDLTTCGPADLMGTPVVGAPLGVYARLAAHGVLEGVNFAATGMLPGCSFIFLLA 180
DB 121 KYIDLTTCGPADLMGTPVVGAPLGVYARLAAHGVLEGVNFAATGMLPGCSFIFLLA 180
QY 121 KYIDLTTCGPADLMGTPVVGAPLGVYARLAAHGVLEGVNFAATGMLPGCSFIFLLA 180
DB 121 KYIDLTTCGPADLMGTPVVGAPLGVYARLAAHGVLEGVNFAATGMLPGCSFIFLLA 180
QY 181 LLSCTITPVSAAEVKIISTGYWNTDCTNDSITWQOAAVLAHVGCVPCCKVNAQCWI 240
DB 181 LLSCTITPVSAAEVKIISTGYWNTDCTNDSITWQOAAVLAHVGCVPCCKVNAQCWI 240
QY 181 LLSCTITPVSAAEVKIISTGYWNTDCTNDSITWQOAAVLAHVGCVPCCKVNAQCWI 240
DB 181 LLSCTITPVSAAEVKIISTGYWNTDCTNDSITWQOAAVLAHVGCVPCCKVNAQCWI 240
QY 241 PVSNNVAVRPGALTOGLRTHIDMVVMSATLCSALVYVDLCGGVMLAQMFIYSPQHMF 300
DB 241 PVSNNVAVRPGALTOGLRTHIDMVVMSATLCSALVYVDLCGGVMLAQMFIYSPQHMF 300
QY 301 VQDCNCSIYPGITIGHRAMDMMMNNSPTATMILAYAMRVEYIIDIIISGHHGVMFGLA 360
DB 301 VQDCNCSIYPGITIGHRAMDMMMNNSPTATMILAYAMRVEYIIDIIISGHHGVMFGLA 360
QY 361 YFSMOGAMAKVVVILLAAVGDARHTVGGSAOQTGRLTSLFDMGPRKIOLVNNGSW 420
DB 361 YFSMOGAMAKVVVILLAAVGDARHTVGGSAOQTGRLTSLFDMGPRKIOLVNNGSW 420
QY 421 HINRTALNCNDSLHTGFIALFYTHSFNSGCEPERMSACRSIEAFVWGALQYEDNVN 480
DB 421 HINRTALNCNDSLHTGFIALFYTHSFNSGCEPERMSACRSIEAFVWGALQYEDNVN 480
QY 481 PEDMRPYCMHYPPRQCGVSAKTVCGPVYCTFSPVYVGGTTDRLAGPYTTGNETDVL 540
DB 481 PEDMRPYCMHYPPRQCGVSAKTVCGPVYCTFSPVYVGGTTDRLAGPYTTGNETDVL 540
QY 541 LNSTRPPLGSMFECTMNNSSGYTKTGAPRCRTRADFNASTDLCTDPCRKHDPDTYK 600
DB 541 LNSTRPPLGSMFECTMNNSSGYTKTGAPRCRTRADFNASTDLCTDPCRKHDPDTYK 600
QY 601 CGSGPWLTPRCLIDYRRLMHPYCTVYTI FKIRMTVGVGVEHRLTAACNFTGRDCNLE 660
DB 601 CGSGPWLTPRCLIDYRRLMHPYCTVYTI FKIRMTVGVGVEHRLTAACNFTGRDCNLE 660
QY 661 RDRSQSLPLHSTTEWAILPCSYSDLPALSTGLLHONVVDYQFMYGSLPATKTYVM 720
DB 661 RDRSQSLPLHSTTEWAILPCSYSDLPALSTGLLHONVVDYQFMYGSLPATKTYVM 720

QY 721 EWWILFLLLADARVACLMMLILGQAEALAEKVLILHAASNASGNGELYFVIFVAAW 780
DB 721 EWWILFLLLADARVACLMMLILGQAEALAEKVLILHAASNASGNGELYFVIFVAAW 780
QY 781 YIKGRVPLATYSLTGLMSFSLILLALPOQAYAYDASVHGQIGALLVMTLFTLPYK 840
DB 781 YIKGRVPLATYSLTGLMSFSLILLALPOQAYAYDASVHGQIGALLVMTLFTLPYK 840
QY 841 TILSRFLMMLCYLLTIGEMVQEWAPMVGRRGDSIIWAAVAFYGVVFDIKMLLA 900
DB 841 TILSRFLMMLCYLLTIGEMVQEWAPMVGRRGDSIIWAAVAFYGVVFDIKMLLA 900
QY 901 GPAYLLKALATVPYFVARAALLRMCTMARHLAAGRYVOMALLALGRMTCTYIDHLTPM 960
DB 901 GPAYLLKALATVPYFVARAALLRMCTMARHLAAGRYVOMALLALGRMTCTYIDHLTPM 960
QY 961 SDMAASGLDLAVAVPIIFSPMEKKVIVWGAETAAQSDILHGLPVSAKREVLIGPAD 1020
DB 961 SDMAASGLDLAVAVPIIFSPMEKKVIVWGAETAAQSDILHGLPVSAKREVLIGPAD 1020
QY 1021 GYTSKMSLAPRTAATAOQTRGLGTVVSMGRDKEQAGEIQLVSTYQSFGLTSISG 1080
DB 1021 GYTSKMSLAPRTAATAOQTRGLGTVVSMGRDKEQAGEIQLVSTYQSFGLTSISG 1080
QY 1081 VMTVYHAGANKTLAAGSRPVTQWSSAEGDLVGMSPSGTSLBECTCGAVDLYLVTRN 1140
DB 1081 VMTVYHAGANKTLAAGSRPVTQWSSAEGDLVGMSPSGTSLBECTCGAVDLYLVTRN 1140
QY 1141 ADVIPARRRQDKRGALLSBRPLSTLKSGSGGVLCPRGHAGVGFRAAVCSRGVAKSIDPI 1200
DB 1141 ADVIPARRRQDKRGALLSBRPLSTLKSGSGGVLCPRGHAGVGFRAAVCSRGVAKSIDPI 1200
QY 1201 PVETLDIYRSPFSDNSNTPPAVPQTYGVYLAHPGSSKSTKVPAAVAAQGVYLVNP 1260
DB 1201 PVETLDIYRSPFSDNSNTPPAVPQTYGVYLAHPGSSKSTKVPAAVAAQGVYLVNP 1260
QY 1261 SVAATLGFAYLSKAHGINPRTGVRVTTGAPITVSYGKFLADGGCAGAYDIIICD 1320
DB 1261 SVAATLGFAYLSKAHGINPRTGVRVTTGAPITVSYGKFLADGGCAGAYDIIICD 1320
QY 1321 ECHAVDSTIILGIVLDQAEFAGVRLVLAATATPPGSYTTPHPIEBVALGOEHIPIY 1380
DB 1321 ECHAVDSTIILGIVLDQAEFAGVRLVLAATATPPGSYTTPHPIEBVALGOEHIPIY 1380
QY 1381 GRAIPLSYIKGRHLLFCHSKKKCELAALGKMGINSVAYYRGLDVSIYPIQGDVVVVA 1440
DB 1381 GRAIPLSYIKGRHLLFCHSKKKCELAALGKMGINSVAYYRGLDVSIYPIQGDVVVVA 1440
QY 1441 TDALMTGYTGFSDVIDCNVAATVQVVDPSLDPTFTTQIVPODAVRSQRGRGTGRRL 1500
DB 1441 TDALMTGYTGFSDVIDCNVAATVQVVDPSLDPTFTTQIVPODAVRSQRGRGTGRRL 1500
QY 1501 GIYRVYSTGERASGMFDSVILCEYDAGAAWELTSEETTVRLRAYFNTPGLEPVCODHLE 1560
DB 1501 GIYRVYSTGERASGMFDSVILCEYDAGAAWELTSEETTVRLRAYFNTPGLEPVCODHLE 1560
QY 1561 FWEAVFTGLTHIDAHFLSOTKSGENFAITVYQAVTCARARAAPPSPVMMVCLTRLRP 1620
DB 1561 FWEAVFTGLTHIDAHFLSOTKSGENFAITVYQAVTCARARAAPPSPVMMVCLTRLRP 1620
QY 1621 TLVGPFPPLLYRGSVYNEVTLLHPVTKYIATCMQADLEWNTSTWLAGVLAAYAYCLA 1680
DB 1621 TLVGPFPPLLYRGSVYNEVTLLHPVTKYIATCMQADLEWNTSTWLAGVLAAYAYCLA 1680
QY 1681 TGCVCIIIGRLHINQRAVVAADXEVLVEAFDEMEBCASRAALIIEGQRIAEMLKSKIQGLL 1740
DB 1681 TGCVCIIIGRLHINQRAVVAADXEVLVEAFDEMEBCASRAALIIEGQRIAEMLKSKIQGLL 1740
QY 1741 QOASROAODIOPVQASMPKVPQFPAKHMNNTISGIQYAGISTLPGNPAVSMMAFSA 1800
DB 1741 QOASROAODIOPVQASMPKVPQFPAKHMNNTISGIQYAGISTLPGNPAVSMMAFSA 1800
QY 1801 LTPSLSTSTIILNLLGWLASQIAPAGATGFVVGIVGAUVGSIQGLKVLVDILAGY 1860

Db 1801 LTSPLESTSTIILNIIIGMGLASQIAPPAGATGFVSGVGAAGVSGIGLGVLTVDIIAGYG 1860
Qy 1861 AGISGALVAFKINGSEKPSMEDVYNLLPGILSPGALVGVYCAILBRHVGPGGAGVQM 1920
Db 1861 AGISGALVAFKINGSEKPSMEDVYNLLPGILSPGALVGVYCAILBRHVGPGGAGVQM 1920
Qy 1921 NRLIAFASRGNHVAPTHYVTESDASQVTOGLSLTTSLLRLHNMITEDECPIPCGGSW 1980
Db 1921 NRLIAFASRGNHVAPTHYVTESDASQVTOGLSLTTSLLRLHNMITEDECPIPCGGSW 1980
Qy 1981 LRVDWVAVCTIILDFFKMWLTSKLPKMPGLPVSQCGYGVWAGTGIMTRCPGAGNIS 2040
Db 1981 LRVDWVAVCTIILDFFKMWLTSKLPKMPGLPVSQCGYGVWAGTGIMTRCPGAGNIS 2040
Qy 2041 GNVALGSMRTITGPKTGNIMOGTEPINCYTRGQCVPPAPAPNFKVALIRVVAASEVAYTOH 2100
Db 2041 GNVALGSMRTITGPKTGNIMOGTEPINCYTRGQCVPPAPAPNFKVALIRVVAASEVAYTOH 2100
Qy 2101 GSYHYITGLTTDNLKVPQOLPSPEFFSVDGVOQIHRPAPTPEKPFDEVSFCVGLNSFV 2160
Db 2101 GSYHYITGLTTDNLKVPQOLPSPEFFSVDGVOQIHRPAPTPEKPFDEVSFCVGLNSFV 2160
Qy 2161 GSQLEPCDPEPPTDVLMSMLTDPSSHITABTAARLARSGSPSEASSASQLSAPSLRATCT 2220
Db 2161 GSQLEPCDPEPPTDVLMSMLTDPSSHITABTAARLARSGSPSEASSASQLSAPSLRATCT 2220
Qy 2221 THGKAYVDVMDVANDLPMGQDVTRIIESSGSKVVLDSLDPMYBERSDLEPSIPESEMLPKKR 2280
Db 2221 THGKAYVDVMDVANDLPMGQDVTRIIESSGSKVVLDSLDPMYBERSDLEPSIPESEMLPKKR 2280
Qy 2281 FPPALPAMARPDYNPPIVSWKRPDYQPAIVAGCALPPPKPTPPPPRRRTVGLSBDSI 2340
Db 2281 FPPALPAMARPDYNPPIVSWKRPDYQPAIVAGCALPPPKPTPPPPRRRTVGLSBDSI 2340
Qy 2341 GDALQOLAISFGQPPPSGDSGLSTGAGADSGSQTPPDLALSTETSISSMPLEBELG 2400
Db 2341 GDALQOLAISFGQPPPSGDSGLSTGAGADSGSQTPPDLALSTETSISSMPLEBELG 2400
Qy 2401 DPDLPEPOVEPQPPQGVAAFGSDSGSMSTCSSEEDSVVCCSMSTYMGALITPPCPPEE 2460
Db 2401 DPDLPEPOVEPQPPQGVAAFGSDSGSMSTCSSEEDSVVCCSMSTYMGALITPPCPPEE 2460
Qy 2461 EKLPIPNLSNLSLRHANKVCTTTKASLSRAKVTFRMOVLDSDYDVLKDIKLAASKV 2520
Db 2461 EKLPIPNLSNLSLRHANKVCTTTKASLSRAKVTFRMOVLDSDYDVLKDIKLAASKV 2520
Qy 2521 TARRLTMBEACQLTPPHSARSKYGFGAKEVASLSGRAVNHIKSVWKDLLEDSETPITTI 2580
Db 2521 TARRLTMBEACQLTPPHSARSKYGFGAKEVASLSGRAVNHIKSVWKDLLEDSETPITTI 2580
Qy 2581 MAKKEVCVDPDTKGKKAARLIIVPDIGVRVCEGALYDITOKLPOAVMGASVGFQSP 2640
Db 2581 MAKKEVCVDPDTKGKKAARLIIVPDIGVRVCEGALYDITOKLPOAVMGASVGFQSP 2640
Qy 2641 QRVFLLKAMAEKKDPMGFSYDTRCFDSTYTERDIRTESIYRACSLPBEAHTAHSLTE 2700
Db 2641 QRVFLLKAMAEKKDPMGFSYDTRCFDSTYTERDIRTESIYRACSLPBEAHTAHSLTE 2700
Qy 2701 RLIVYGEMFNSKQTCGRRCRASGULTTSMGNTITCYVALAACAKAGIIAPMLVCGD 2760
Db 2701 RLIVYGEMFNSKQTCGRRCRASGULTTSMGNTITCYVALAACAKAGIIAPMLVCGD 2760
Qy 2761 DLVVISISQGTBEEDERNLRAFTEAMTRYAPGPDPPEPEYDLIELITSCSSNVASALDPOG 2820
Db 2761 DLVVISISQGTBEEDERNLRAFTEAMTRYAPGPDPPEPEYDLIELITSCSSNVASALDPOG 2820
Qy 2821 RRRYYLTRDPTTPIARAAMETVRHSVNVSMGNIIQYAPTIMAMVMTLHFFSILMAQDT 2880
Db 2821 RRRYYLTRDPTTPIARAAMETVRHSVNVSMGNIIQYAPTIMAMVMTLHFFSILMAQDT 2880
Qy 2881 LDQNLNRMGAUVSVSPDLPAIIBRLHGIDASLTHTYTHELTRVASALRKUGAPPLR 2940
Db 2881 LDQNLNRMGAUVSVSPDLPAIIBRLHGIDASLTHTYTHELTRVASALRKUGAPPLR 2940

Db 2881 LDQNLNRMGAUVSVSPDLPAIIBRLHGIDASLTHTYTHELTRVASALRKUGAPPLR 2940
RESULT 6
ID ABG30688 standard; protein; 3033 AA.
XX
XX ABG30688;
XX
XX 07-OCT-2002 (first entry)
XX
XX Human HCV-related polypeptide.
XX
XX Hepatitis C virus; human; virucide; gene therapy; HCV;
XX
XX fulminant hepatitis C.
XX
XX Homo sapiens.
XX
XX JP2002171978-A.
XX
XX 18-JUN-2002.
XX
XX 01-DEC-2000; 2000JP-00367365.
XX
XX 01-DEC-2000; 2000JP-00367365.
XX
XX 01-DEC-2000; 2000JP-00367365.
XX
XX (TOKR-) ZH TOKYOTO RINSHO IGAKU SOGO KENKYUSHO.
XX
XX (TORA) TORAY IND INC.
XX
XX WPI, 2002-569884/61.
XX
XX N-PSDB; ABR88904.
XX
XX A gene of a fulminant hepatitis C virus strain and the encoded
XX polypeptide useful in gene therapy to treat hepatitis C.
XX
XX Claim 1; Page 25-33; 36pp; Japanese.
XX
XX
XX The invention relates to a human polypeptide related to hepatitis C virus
XX (HCV), and the polynucleotide encoding it. The polypeptide can be used
XX for the development of gene therapy on fulminant hepatitis C. This
XX sequence represents a human HCV-related polypeptide
XX
XX
XX Sequence 3033 AA;
Query Match 92.7%; Score 14943; DB 5; Length 3033;
Best Local Similarity 91.2%; Pred. No. 0;
Matches 2765; Conservative 129; Mismatches 139; Indels 0; Gaps 0;
Qy 1 MSTNPKPQKRTKRNTRNPQDVKPPGQGI VGGVYLLPRGRPLGVRATRKTSERSQPRG 60
Db 1 MSTNPKPQKRTKRNTRNPQDVKPPGQGI VGGVYLLPRGRPLGVRATRKTSERSQPRG 60
Qy 61 RROPPIKDRRSKSGSKGKRPMPLYGNGLCMAGMLSPRSRPSWGBNDPRHSRNVG 120
Db 61 RROPPIKDRRSKSGSKGKRPMPLYGNGLCMAGMLSPRSRPSWGBNDPRHSRNVG 120
Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGGVABALAHGVYVLEDDGVNFATGNLPGCSFIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGGVABALAHGVYVLEDDGVNFATGNLPGCSFIFLLA 180
Qy 181 LLSCTITPVSAABVKNISITGVYVNTDCTNDSITWQLOAAVLHVPGCVPCERKYGNASQCM 240
Db 181 LLSCTITPVSAABVKNISITGVYVNTDCTNDSITWQLOAAVLHVPGCVPCERKYGNASQCM 240
Qy 241 PVSNNVAVRQPALVQGLTHIDMVVMSATFCSALYVGLCGVMTAAAVFTVSPQYHWF 300
Db 241 PVSNNVAVRQPALVQGLTHIDMVVMSATFCSALYVGLCGVMTAAAVFTVSPQYHWF 300
Qy 301 VQDCNCSIIPGTTGRRMAMDMMMNMSPTATMILAVAMVPRVIIDIIISGAHMGVFGDLA 360
Db 301 VQDCNCSIIPGTTGRRMAMDMMMNMSPTATMILAVAMVPRVIIDIIISGAHMGVFGDLA 360
Qy 361 YFSMGAAMAKVVVILLIAGVDAHTHTVGSAAQTGRLTSLFDMGPRQKIQLVNTNSGW 420
Db 361 YFSMGAAMAKVVVILLIAGVDAHTHTVGSAAQTGRLTSLFDMGPRQKIQLVNTNSGW 420

```

Db      361 YFSMQAMAKVIYILLAAAGVADAGTTVGAVARSINVLAGVFSHPQONIQIINTNGSM 420
Qy      421 HINRPLANCNDLSHTGFIASLPTTHSPNSSGCEBRSACRSIAFVNGALQYEDNVTR 480
Db      421 HINRPLANCNDLSHTGFIASLPTTHSPNSSGCEBRSACRSIAFVNGALQYEDNVTR 480
Qy      481 PEDMRPYCHNYPPOCGVYSAKTVGAPVCFTPSPVYVGTDTDLGAPTYWGENETDVEI 540
Db      481 PEDMRPYCHNYPPOCGVYSAKTVGAPVCFTPSPVYVGTDTDLGAPTYWGENETDVEI 540
Qy      541 LNSTRPPLGSMWEGCTMWNSSGYTKTCGAPPCRTRADFNASTDLCEPTDFRKAPDTTYLK 600
Db      541 LNSTRPPLGSMWEGCTMWNSSGYTKTCGAPPCRTRADFNASTDLCEPTDFRKAPDTTYLK 600
Qy      601 CGSGPMLTRCLIDYRRLMHPYCTVNTYTFKIRMYGVGVYERHLTAACNFTGRDCNLED 660
Db      601 CGSGPMLTRCLIDYRRLMHPYCTVNTYTFKIRMYGVGVYERHLTAACNFTGRDCNLED 660
Qy      661 RDRSOLPLHSTTEWAILPCSYSDLPALSTGLHLHONIVDYQFMYGSPALTKYIVRM 720
Db      661 RDRSOLPLHSTTEWAILPCSYSDLPALSTGLHLHONIVDYQFMYGSPALTKYIVRM 720
Qy      721 EWWILLFLLADARVACIMMLILGQAEALAEKVLHAASNASCNGLYFVIFPVAAM 780
Db      721 EWWILLFLLADARVACIMMLILGQAEALAEKVLHAASNASCNGLYFVIFPVAAM 780
Qy      781 YIKGRVPLATYSLTGLASFSLILLAPQOAYAYDASVHGQIGALLWMTLFTLTPTGK 840
Db      781 YIKGRVPLATYSLTGLASFSLILLAPQOAYAYDASVHGQIGALLWMTLFTLTPTGK 840
Qy      841 TLLSRFLMWCYLLTLGEBAMVOEMAPPMQVGRGRDAIIVAVAI FYRPGVFDITKMLAVL 900
Db      841 TLLSRFLMWCYLLTLGEBAMVOEMAPPMQVGRGRDAIIVAVAI FYRPGVFDITKMLAVL 900
Qy      901 GBAYILLKGLATRVYFVRAHALLRMCTMARHLAGRYVOMALLALGRMTGYIYDHLTBM 960
Db      901 GBAYILLKGLATRVYFVRAHALLRMCTMARHLAGRYVOMALLALGRMTGYIYDHLTBM 960
Qy      961 SOWMAASGLDLAVANEPILFSPMEKKVIYWGAEFTAACGDIHLGLPVSAFLGRBVLGPD 1020
Db      961 SOWMAASGLDLAVANEPILFSPMEKKVIYWGAEFTAACGDIHLGLPVSAFLGRBVLGPD 1020
Qy      1021 GYTSKMSLLAPITLTAOQTRGLGTIVSMTRGRDTEOAGEIOVLSTYOSFLGTSISG 1080
Db      1021 GYTSKMSLLAPITLTAOQTRGLGTIVSMTRGRDTEOAGEIOVLSTYOSFLGTSISG 1080
Qy      1081 VLMTVYHGAAGNTLAGSRGPVTQWYSSAEGDLVGMPSPEGTSLBECTCGAVDLYLVTBN 1140
Db      1081 VLMTVYHGAAGNTLAGSRGPVTQWYSSAEGDLVGMPSPEGTSLBECTCGAVDLYLVTBN 1140
Qy      1141 ADVIPARRRGDKRGALLSPRPISTLKSSGGPYLCRGAHGVFRAVCSRGVANSIDPI 1200
Db      1141 ADVIPARRRGDKRGALLSPRPISTLKSSGGPYLCRGAHGVFRAVCSRGVANSIDPI 1200
Qy      1201 PVETIDIVTRSPFSPNSNTPPAVPOTYQYGYLHAPFGSGSKTKVPAYAAOGKVLVLP 1260
Db      1201 PVETIDIVTRSPFSPNSNTPPAVPOTYQYGYLHAPFGSGSKTKVPAYAAOGKVLVLP 1260
Qy      1261 SYAATLFGFAYLSKAHGINPNIRTVGRVTTGAPIITYSTYKFLADGGCAGAYDIIICD 1320
Db      1261 SYAATLFGFAYLSKAHGINPNIRTVGRVTTGAPIITYSTYKFLADGGCAGAYDIIICD 1320
Qy      1321 ECHAVDSTIILGITYLDOAETNAGVRLTVLATRTPPGSATTTPHPIEVALGOEGEIPRY 1380
Db      1321 ECHAVDSTIILGITYLDOAETNAGVRLTVLATRTPPGSATTTPHPIEVALGOEGEIPRY 1380
Qy      1381 GRAIPLSYIKGGSHLIFCHSKKKCDELAALRGMGNSVAYYRGDLVSYIPIPOGDDVVVA 1440
Db      1381 GRAIPLSYIKGGSHLIFCHSKKKCDELAALRGMGNSVAYYRGDLVSYIPIPOGDDVVVA 1440
Qy      1441 TDALMTGYTGFDSVIDCNVAVTQVDFSLDPTFTITQIVPODASRSQRRGRTGRGL 1500

```

```

Db      1441 TDALMTGYTGFDSVIDCNVAVTQVDFSLDPTFTITQIVPODASRSQRRGRTGRGL 1500
Qy      1501 GYRVSUTGERASGMPDSVYLCEVDAGAAWELPSETTVLRAVFNTPGLPYCODHLE 1560
Db      1501 GYRVSUTGERASGMPDSVYLCEVDAGAAWELPSETTVLRAVFNTPGLPYCODHLE 1560
Qy      1561 FWEAVFTGLTHIDAFPLSOTKSGENFAYLTAYOATVCARALAPPSMDVMKCLTRLKP 1620
Db      1561 FWEAVFTGLTHIDAFPLSOTKSGENFAYLTAYOATVCARALAPPSMDVMKCLTRLKP 1620
Qy      1621 TLVGPTPLLRLGSVTEVTLTHPVTKYIATCMQADLEVMSTTWVLAAVAAVYCLIA 1680
Db      1621 TLVGPTPLLRLGSVTEVTLTHPVTKYIATCMQADLEVMSTTWVLAAVAAVYCLIA 1680
Qy      1681 TGCVCIGSLHINORAVVAPDXEVLVEAPDEMECASRALLEBQRIAEMLKSKIOGLL 1740
Db      1681 TGCVCIGSLHINORAVVAPDXEVLVEAPDEMECASRALLEBQRIAEMLKSKIOGLL 1740
Qy      1741 QOASQOADIOPVQOASMPKVEQFNAKHMNFISGIOYLAGLSTLPGNPAVASMAFSA 1800
Db      1741 QOASQOADIOPVQOASMPKVEQFNAKHMNFISGIOYLAGLSTLPGNPAVASMAFSA 1800
Qy      1801 LTPSLSTSTILLNITLGGMLASQIAPAGATGFVVSGLVGAAVGSIIGLKVLDLIAYG 1860
Db      1801 LTPSLSTSTILLNITLGGMLASQIAPAGATGFVVSGLVGAAVGSIIGLKVLDLIAYG 1860
Qy      1861 AGISGALVAFKIMSEKPEMEDVNLPGIISPGALVGVICAILLRHVGBEGAVQM 1920
Db      1861 AGISGALVAFKIMSEKPEMEDVNLPGIISPGALVGVICAILLRHVGBEGAVQM 1920
Qy      1921 NRIIPASRGNVAFTHYTESDASQRTVQLGSLTISLRLHNMITDEDEPIPCGSGM 1980
Db      1921 NRIIPASRGNVAFTHYTESDASQRTVQLGSLTISLRLHNMITDEDEPIPCGSGM 1980
Qy      1981 IRDVMWVCTIITDFRKNMITSKLPRKMGLPFVSCQKGYKGWAGTGMTRCCGANIS 2040
Db      1981 IRDVMWVCTIITDFRKNMITSKLPRKMGLPFVSCQKGYKGWAGTGMTRCCGANIS 2040
Qy      2041 GNVRLGSRNITGPKTCMNIWQCTPPINCTEBOCPKRAPNKAIVRAASEVAVTQH 2100
Db      2041 GNVRLGSRNITGPKTCMNIWQCTPPINCTEBOCPKRAPNKAIVRAASEVAVTQH 2100
Qy      2101 GSXYITGLTNDNLKVPCLPSEFFSWDGVQIRPAPTPPFRDEVSFCVGLNSPVV 2160
Db      2101 GSXYITGLTNDNLKVPCLPSEFFSWDGVQIRPAPTPPFRDEVSFCVGLNSPVV 2160
Qy      2161 GSQLPCEBPDVDVLSMLTDBSHITTAETARLARSGPSSBASASQASAPSLRATCT 2220
Db      2161 GSQLPCEBPDVDVLSMLTDBSHITTAETARLARSGPSSBASASQASAPSLRATCT 2220
Qy      2221 THSNITVDVMDANILMEGVAQTEPESRVPPLDLLEPAABESDLBPSIPEECMLPRSG 2280
Db      2221 THSNITVDVMDANILMEGVAQTEPESRVPPLDLLEPAABESDLBPSIPEECMLPRSG 2280
Qy      2281 FPPALPAMARPYNPLVESWKRPDYQATVAGCALPPEPKPTPPPPRRRTVGLSEDSI 2340
Db      2281 FPPALPAMARPYNPLVESWKRPDYQATVAGCALPPEPKPTPPPPRRRTVGLSEDSI 2340
Qy      2341 GDALQOLAIKSFQPPSGDSGLSTGAGAADSQTPDEDLALSETGSISSMPLEGEIG 2400
Db      2341 GDALQOLAIKSFQPPSGDSGLSTGAGAADSQTPDEDLALSETGSISSMPLEGEIG 2400
Qy      2401 DDLLEPBOVPOPPQGVAAAGSGSGSWSTSEEDDVTVCSSMSYSWGTALITPCSPE 2460
Db      2401 DDLLEPBOVPOPPQGVAAAGSGSGSWSTSEEDDVTVCSSMSYSWGTALITPCSPE 2460
Qy      2461 EKLPIINPLSNSILRYHNKYCTTTSKASLRAKAVTFDMQVLDYSYDSVLKDIKLAASKV 2520
Db      2461 EKLPIINPLSNSILRYHNKYCTTTSKASLRAKAVTFDMQVLDYSYDSVLKDIKLAASKV 2520
Qy      2521 TABLLTMBACOLTPPHSARSKYGFALVNSLSGRAVNHKSVKMDLLEDSPTPIPTTI 2580
Db      2521 TABLLTMBACOLTPPHSARSKYGFALVNSLSGRAVNHKSVKMDLLEDSPTPIPTTI 2580

```

QY 2581 MANKVFCVDPPTGKGAARLIYVPLGVVCEKMAIYDITQKLPQAVMGASYGFOYSPA 2640
Db 2581 MANKVFCVDPPTGKGAARLIYVPLGVVCEKMAIYDITQKLPQAVMGASYGFOYSPA 2640
QY 2641 QRYEFLIKAKAEKKDPMGFSYDTRCPDSTYTERDINTESITYACSLPERAHTAHSLTE 2700
Db 2641 QRYEFLIKAKAEKKDPMGFSYDTRCPDSTYTERDINTESITYACSLPERAHTAHSLTE 2700
QY 2701 RLTVGGPMFNSKQOTGCRRCRASGVLTSMGNTITCYVALAACRAAGIAPFTMLVCPD 2760
Db 2701 RLTVGGPMFNSKQOTGCRRCRASGVLTSMGNTITCYVALAACRAAGIAPFTMLVCPD 2760
QY 2761 DLVVISBSQGTEDERNLRAFTAMTRYSAPOGPDPREYDELITSCSSNVVALGPGQ 2820
Db 2761 DLVVISBSQGTEDERNLRAFTAMTRYSAPOGPDPREYDELITSCSSNVVALGPGQ 2820
QY 2821 RRRYVLTRODPTTPLARAAMETVSHSPNSWLNIGIYAPPTIWMVLMTHFFSILMVDOT 2880
Db 2821 RRRYVLTRODPTTPLARAAMETVSHSPNSWLNIGIYAPPTIWMVLMTHFFSILMVDOT 2880
QY 2881 LDONLNFEMYGAVYSVPLDLPALIERLHGLDAFSLHTYTPHELTVAASLRDLGAPLR 2940
Db 2881 LDONLNFEMYGAVYSVPLDLPALIERLHGLDAFSLHTYTPHELTVAASLRDLGAPLR 2940
QY 2941 AMKSRARAVASLISRGRAAVCGRYLFNNAVTKLKLPLPEARLLDLSMFTVAGAGG 3000
Db 2941 VMSRARAASLISRGKAAVCGRYLFNNAVTKLKLPLPEARLLDLSMFTVAGAGG 3000
QY 3001 DIYHSVSRAPRLILFGLLILFVGVGLFLPAR 3033
Db 3001 DIYHSVSRAPRLILFGLLILFVGVGLFLPAR 3033

RESULT 7

AAR33539 standard; protein; 3033 AA.

AC AAR33539;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 01-JUL-1993 (first entry)
XX
DE NANBH virus strain HC-J8 protein.
XX
KM Non A non B hepatitis virus; plasma; degenerate.
XX
OS Non-A.
OS non-B hepatitis virus.
PN EP532167-A2.
XX
PD 17-MAR-1993.
XX
PF 30-JUL-1992; 92EP-00306952.
XX
PR 09-AUG-1991; 91JP-00287402.
PR 05-DEC-1991; 91JP-00360441.
XX
PA (IMMO) IMMUNO JAPAN INC.
XX
PI Okamoto H, Nakamura T;
XX WPI; 1993-087166/11.
DR P-PSDB; AAQ38221.
XX
PT Polynucleotide(s), polypeptide(s) and antibodies of NANBH virus - useful
XX for detecting NANBH, as a vaccine and for screening blood samples.
PS Claim 8; Page 65-79; 93pp; English.
XX
CC RNA was isolated from the plasma of human patients positive for NANBH

CC virus (strain HC-J8) and was subjected to reverse transcription to
CC produce cDNA. The resulting cDNA was amplified by PCR, and nucleic acid
CC sequences determined by analysis of clones obd. by PCR amplification (42
CC clones in total). The NANBH HC-J8 genome was found to contain a
CC degenerate open reading frame encoding polypeptide precursors of 3033
CC amino acid residues. See also AAR33538 and AAR33214. (Updated on 25-MAR-
CC 2003 to correct FN field.) (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 3033 AA;

Query Match 86.5%; Score 13944; DB 2; Length 3033;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 2540; Conservative 247; Mismatches 246; Indels 0; Gaps 0;

QY 1 MSTNPKPQRTKRNTRRRPQDVKFPQGQIYGVYLLRRRGRLGRATRTKTSRQPRG 60
Db 1 MSTNPKPQRTKRNTRRRPQDVKFPQGQIYGVYLLRRRGRLGRATRTKTSRQPRG 60
QY 61 RRQPIPKDRRSTGKSGKPGYPMPLYGNBGCMGLSPRGSRPSPGPNDRHRSRVNG 120
Db 61 RRQPIPKDRRSTGKSGKPGYPMPLYGNBGCMGLSPRGSRPSPGPNDRHRSRVNG 120
QY 121 KTIIDLTCGFADLMGYIVNGAPPLGAVARALAHGRVLEDCVNFATGMLPGCSFSIFLLA 180
Db 121 KTIIDLTCGFADLMGYIVNGAPPLGAVARALAHGRVLEDCVNFATGMLPGCSFSIFLLA 180
QY 122 RVIDITITGFAIDLMGYIVNGAPVGVARALAHGRVLEDCVNFATGMLPGCSFSIFLLA 180
Db 122 RVIDITITGFAIDLMGYIVNGAPVGVARALAHGRVLEDCVNFATGMLPGCSFSIFLLA 180
QY 181 ILSCITTPSAAEVNGNISGVYVNTDCTNDSITWQLQAAVLHPGCVCEKGNASQCM 240
Db 181 ILSCITTPSAAEVNGNISGVYVNTDCTNDSITWQLQAAVLHPGCVCEKGNASQCM 240
QY 241 PVSNNVAVORPGALTOGLRTHIDMVMSATLCSALYVDLCGVMLAQMPLVSPQHMF 300
Db 241 PVSNNVAVORPGALTOGLRTHIDMVMSATLCSALYVDLCGVMLAQMPLVSPQHMF 300
QY 241 QVTNNVAVHGRALTRSLRSHVDMVMAATCASALYVDVCGAVMILSQAFVNSPQRHNF 300
Db 241 QVTNNVAVHGRALTRSLRSHVDMVMAATCASALYVDVCGAVMILSQAFVNSPQRHNF 300
QY 301 VDDCNCSITPGTITGHRAMDMNMNMSPATMILYARVRPVIIDISGAHWGMFGLA 360
Db 301 VDDCNCSITPGTITGHRAMDMNMNMSPATMILYARVRPVIIDISGAHWGMFGLA 360
QY 361 YFSMOGAAKAVVILLIAGVADARTHTVGSAAOTTGRLTSFLDMPGPROKTQIVNTNSGM 420
Db 361 YFSMOGAAKAVVILLIAGVADARTHTVGSAAOTTGRLTSFLDMPGPROKTQIVNTNSGM 420
QY 421 HINRTALNCNDSLATGFIASLFTYTHSPNSGCEPRMSACRSIEARVGMALQYEDNVTN 480
Db 421 HINRTALNCNDSLATGFIASLFTYTHSPNSGCEPRMSACRSIEARVGMALQYEDNVTN 480
QY 481 PEDMRPYCHHPYPRROGVASAKTVGPGVYCFPPSPVVNGTTDRGLAPRYTGENSETDVFL 540
Db 481 PEDMRPYCHHPYPRROGVASAKTVGPGVYCFPPSPVVNGTTDRGLAPRYTGENSETDVFL 540
QY 541 INSTRPPLGSMFGCTMMNSSGYTKTCGAPPCRTADFNASTDULCPTCCFRKHPTTYLK 600
Db 541 INSTRPPLGSMFGCTMMNSSGYTKTCGAPPCRTADFNASTDULCPTCCFRKHPTTYLK 600
QY 601 CGAGPMLTFRCLVDYPRILMHPCTVNFETIFARMYGVGEVHRFSAACNFTGDRCRLED 660
Db 601 CGAGPMLTFRCLVDYPRILMHPCTVNFETIFARMYGVGEVHRFSAACNFTGDRCRLED 660
QY 661 RDRSQSLPLHSTTETMALIPCSYSDLPALSTGLLHQNIVDVQFMGSLPTLYIYRW 720
Db 661 RDRSQSLPLHSTTETMALIPCSYSDLPALSTGLLHQNIVDVQFMGSLPTLYIYRW 720
QY 721 EWWILLFLLADARVACILMLILIGQAAALAEKLVIIHSAASACNGFLYEVIFVAVM 780
Db 721 EWWILLFLLADARVACILMLILIGQAAALAEKLVIIHSAASACNGFLYEVIFVAVM 780
QY 781 YIKGRVPLATYSLTGLMSFSLILLALPQAAVAYDASVHQIGALLVNTLFTLTPGYK 840
Db 781 YIKGRVPLATYSLTGLMSFSLILLALPQAAVAYDASVHQIGALLVNTLFTLTPGYK 840
QY 841 TLISRFMLWLCYLLTGERAMVBNAPPMQVGRGRCITIAVAIFPGVAFDITKLLAVL 900
Db 841 TLISRFMLWLCYLLTGERAMVBNAPPMQVGRGRCITIAVAIFPGVAFDITKLLAVL 900

Qy	901	GPALLLGGALTRVYVYFRAHALLMCMCMABHLLAGGRVYVQAMLLAGMTGTYIYDTHLTPM	960
Dp	901	GPALLLQASLSLRITYFVARAHALLVCTLVKRLAGARYIOMLLITTGMTGYIYDHLSP	960
Qy	961	SDMAASGLRDLAVAVEPIIFSPMEKUYIWMCAETPAACGDILHGLPVBARLGREVLGLPAD	1020
Dp	961	STMAAOGRLDLAIVAEPVFSPMEKUYIWMCAETPAACGDILHGLPVBARLGREVLGLPAD	1020
Qy	1021	GYTSKMSLPLATAYAAQTRFGLLGTIVSMTGDKTEQAGEIOVLSVTYQSFLGTSISG	1080
Dp	1021	GYTSKMWMLLAPITAYYQOTRGLLGAIVSLDTGDKTEQAOQOVLSVTQTFLGTSISG	1080
Qy	1081	VLMTVYHAGAKKTLAGSGPYTQWYSSAEGELVCMPSPPGKSLDEPTCCGAVDLYLVTBN	1140
Dp	1081	VLMTVYHAGAKKTIAGPKGPYTQWYTSAGELVCMPSPPGKSLDEPTCCGAVDLYLVTBN	1140
Qy	1141	ADVIPARRRDKRGALLSPRLSTLKGSSGGPYLCPRHAGVFERAAVCSRGVAKSIDFI	1200
Dp	1141	ADVIPRRKDRRALLSPRLSTLKGSSGGPYLCSSGHAGLPPAAVCAARGVAKSIDFI	1200
Qy	1201	PVELLDIYTRSPTSGSDNSTPPAVQOTQOVGLHPTSGSKTKPVVYAAQGYVUYLNP	1260
Dp	1201	PVESLDVATRPSPSDNSTPPAVQOSTQOVGLHPTSGSKTKPVVYASQGYVUYLNP	1260
Qy	1261	SVAATTLGFGALSLSAHGINPNIIRIGVPTVTTCABITYSTGKFLADGGCAGAYDIICD	1320
Dp	1261	SVAATTLGFGAAMSQAHGINPNIIRIGVPTVTTCGDSITYSTGKFLADGGKAAGAYDIICD	1320
Qy	1321	ECHAVDSTTLTIGITVLDOAETAGVRLTVLATATPPQSVTTPHPNIEBVALGOEGEIPFY	1380
Dp	1321	ECHSVDAATTLTIGITVLDOAETAGVRLTVLATATPPGVVTPPHNIEBVALGHEGGEIPFY	1380
Qy	1381	GRALPLASYIKGGRHLIFCHSKKKCDLALALRGKMSVAAAYRGLDVSVPQTQDVVVVA	1440
Dp	1381	GKALPLAFIKGGRHLIFCHSKKKCDLALALRGKGVNAVAAYRGLDVSVPQTQDVVVVA	1440
Qy	1441	TDALMTGYTGDFDSVINDCNAAVTVQVDFSLDPTFTITQYIVPODAVRSQGRGRTGRGL	1500
Dp	1441	TDALMTGYTGDFDSVINDCNAAVSQIVDFSLDPTFTITQYIVPODAVRSQGRGRTGRGL	1500
Qy	1501	GIYVYVYSGEASGSMFDSVYLCCECYDAGAAAYELTBSSETVYRLAAYPNTQGLPYCDHLE	1560
Dp	1501	GVYVYVYSGGEPSSGMPFDSVYLCCECYDAGAAAYELTLPATVYRLAAYPNTQGLPYCDHLE	1560
Qy	1561	FMEAVFTGLTRI DAHPLFSQTKQSGENFAYLAAOYATCARAKAPPPSMDVMWKCLTRLPK	1620
Dp	1561	FMEAVFTGLTHIIDAHPLFSQTKQSGENFAYLYATYATCARAKAPPPSMDVMWKCLTRLPK	1620
Qy	1621	TLVGPPTPLLYRLSGVTNEVTLTHPVTXYIATCMQADLEBWTSTWVLAGVLAAYAAVCLA	1680
Dp	1621	TLVGPPTPLLYRLGAVTNEVTLTHPVTXYIATCMQADLEIMTSSWVLAGVLAAYAAVCLA	1680
Qy	1681	TGCACCIIGRLHINORAVAVAPPKEXLYLEAFDEMECASPAALIEERGRIAEMLKSKIOGL	1740
Dp	1681	TGCISIIIGRLHINORAVAVAPPKEXLYLEAFDEMECASPAALIEERGOMAEMLKSKIOGL	1740
Qy	1741	QOASQKQADIDPTVOASWPKVEQEFWAKHMNFISGIQYLAGLSTLPGNPAVASMAFSA	1800
Dp	1741	QOATROADIDIPAIQSSWPKLEQEFWAKHMNFISGIQYLAGLSTLPGNPAVASMAFSA	1800
Qy	1801	LTSPLSTSTTLINLTIGMGLASQIAPAPGAGTFVYSGLYGAAVYSIGLKYLVNLIAGYG	1860
Dp	1801	LTSPLPSTSTTLINIMGMGLASQIAPAPGAGTFVYSGLYGAAVYSIGLKYLVNLIAGYG	1860
Qy	1861	AGISGALVAFKIMSGEKSPMEDVYNLPGJLSPEALVGVYCAIILRRHYPGEGAYQM	1920
Dp	1861	AGISGALVAFKIMSGEKPTVEDVYNLPAIISPALVGVYCAIILRRHYQSGEGAYQM	1920
Qy	1921	NRLIAFASRGHVAPTHYTESDASQRYTQLLSGLTTSLRLRLHMTEDCPIPCCGSW	1980
Dp	1921	NRLIAFASRGHVAPTHYVESDASQRYTQVLSLTTSLRLRLHMTEDCCPPCCGSW	1980

QY	1981	LRBVMWVCGTILTDFNGWLTSLFPMBPLPVSQCKQYKGVMAGTGIMTRCCGANIS	2040
Db	1981	LQIIMDWCSILTDFKMLSSKULFRMBPPIITSCQKGYKGMWAGTGMTRCCGANIS	2040
QY	2041	GNVRLGSMRITGBKTCMNIWGTFFPINCYTEGCQVPKPAFNFKVAIWRVAEYAVTOH	2100
Db	2041	GHWRMGTMKITGBKTCMNIWGTFFPINCYTEGCPVKPPPNVKTAIWRVAEYAVTOH	2100
QY	2101	GSVHYITGLTTONLKYPCOLPBPFPWSWDGQIHRFAPTRKPPRRDVSFCUGNSRV	2160
Db	2101	GSFSYVTGLTSDULKVPCQVPAPFPFSWVDGQIHRFAPVQCFPRDRVTVTGLNSRV	2160
QY	2161	GSOLPCDPEBDTVMLSMTLDSHSHTAETAAARLARGSPSASSASSQLSAPSLRATCT	2220
Db	2161	GSOLPCDPEBDTVMLSMTLDSHSHTAETAAARLARGSPSQASSASSQLSAPSLKATCT	2220
QY	2221	THGKAYDVMVDANLPMGSDVTRIEBGSKVVLVDSLDPMBEBSDLPSITSEYMLPKR	2280
Db	2221	THRTAADCDMVDANLPMGSDVTRIEBDSKVITLSDLSMTBEVDBDRPFSVSEYLIKRR	2280
QY	2281	FPALPAMARPDVNPPLVBSWKRPPYQAPRVYVGCALPBRKTPPPRPPRRRTVGSSEDSI	2340
Db	2281	FPALPAPMARPDVNPPLVITMKRPGYEPTVUGCALPRTQTPPPRPPRRRAKVLTDQNV	2340
QY	2341	GDAIQQALIKSFQGPSPSGDSLSTGAGAADSQTPPDELALSETGSISSMPLEBELG	2400
Db	2341	EGVLREMADKVLSPLDNNDSGHSHTGADTCGDIYQGPDEFETASBAGLSISMPLEBPG	2400
QY	2401	DPDLBEEQVYEPQPPPGQYVLAAPGSDSGSSTCSSEBDSVCCSMSGYMTGALITPPCSBEE	2460
Db	2401	DPDLFEBPVGASAPPSBEGCEVIDSOSKMSYTVDSDESDVICCSMSGYMTGALITPPCGEE	2460
QY	2461	EKPIPIPLNSILRKYNYKYCTTTSASILRAKAKVFPDMQVLDYSYDVSVDLKDILKLAASKV	2520
Db	2461	EKPIPIPLNSILMRFNKKYSTTSSASILRAKAKVFPDVOVLDAHYSVDLQDVBRASKV	2520
QY	2521	TABULTMEBAQCTTPHSHASRYGFGAKEVNSLSGRAVNIHKSVMKDLLEDSEPIPTTI	2580
Db	2521	SARLLTVEEACALTPHSHAKSRYGFGAKEVNSLSRAVNIHRSWEDLLBEDHEPTIDTTI	2580
QY	2581	MARNEVFCVDPMTGKGKKARLIVYDILGRVCEKALVDITQKLPQAVMGASVGFQYSPA	2640
Db	2581	MARNEVFCIDTPYGGKKAPARLIVYDILGRVCEKALVDIQKLPKALMGFSYGFQYSPA	2640
QY	2641	QRYEFLIKMAWBKKDPMGFSYDTRCFDSTVTERDITRETSIYRASCSPREAHTAIHSITE	2700
Db	2641	ERYDFFILKXWAGSKDDPMGFSYDTRCFDSTVTERDITRETSIYQACSLPQEARTVIHSITE	2700
QY	2701	RLVYGGPMRPSKQCTGGRCRASGVLTTSKMTNTTCYVAKLAAACKAAGIIAPMLVCGD	2760
Db	2701	RLVYGGPMRNSKQSGYRCCRASGFTTSMKMTNTCYIKLAAACKAAGIADVPMVLVCGD	2760
QY	2761	DLVVISQGTDEDEENLRAFTAEAMTRRSAPAGDPRPREYVLELITSCSSNVSAVALGPOG	2820
Db	2761	DLVVISQGNEDENLRAFTAEAMTRRSAPAGDPRPREYVLELITSCSSNVSAVALDSRG	2820
QY	2821	RRRYVLTRODPTPIARANETVRRSPVNSWLGNIIOYAPTTWARMVLTMTFFSLMAODT	2880
Db	2821	RRRYFLTRDPTTPIAAMETVRRSPVNSWLGNIIOYAPTTWARMVIMTFFSILLADOT	2880
QY	2881	LQDNLMFNWYGAVYSVPLDPAIIERLHGDAFSLHYTTHBELTRVSAARLKUGAPLR	2940
Db	2881	LQDNLMFNEMGAVYSVNPDLPAIIERLHGDAFSLHYTSPHELSRVAATTRKUGAPLR	2940
QY	2941	AMGSRARAVASLISRGRAAVCGEYILFNMAVKTLLKUTPLPEARLIDLSSMFTVGAQGG	3000
Db	2941	AMGSRARAVASLIAQGAARALCGEYILFNMAVKTLLKUTPLPEARSLDLSGMFTVGAQGG	3000
QY	3001	DIYHSVSRARPRILLFGLLLFVVGGLFLPLPAR	3033
Db	3001	DIYHSVSHARPRILLCLLLLSVGIGIFLLPLPAR	3033

RESULT 8
 AAR33214
 ID AAR33214 standard; protein; 3033 AA.
 XX
 AC AAR33214:
 XX
 DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 01-JUL-1993 (first entry)
 XX
 DE NANBH virus strain HC-J8 protein.
 XX
 KM Non A non B hepatitis virus; plasma; degenerate.
 XX
 OS Non-A.
 OS non-B hepatitis virus.
 XX
 PN EP532167-A2.
 PD 17-MAR-1993.
 XX
 PF 30-JUL-1992; 92EP-00306952.
 XX
 PR 09-AUG-1991; 91JP-00287402.
 PR 05-DEC-1991; 91JP-00360441.
 XX
 PA (IMMO) IMMUNO JAPAN INC.
 XX
 PI Okamoto H, Nakamura T;
 XX
 DR WPI, 1993-087166/11.
 DR P-PSDB; AAQ38221.
 XX
 PT Polynucleotide(s), polypeptide(s) and antibodies of NANBH virus - useful
 PT for detecting NANBH, as a vaccine and for screening blood samples.
 XX
 PS Claim 9; Page 80-93; 93pp; English.
 XX
 CC RNA was isolated from the plasma of human patients positive for NANBH
 CC virus (strain HC-J8) and was subjected to reverse transcription to
 CC produce cDNA. The resulting cDNA was amplified by PCR, and nucleic acid
 CC sequences determined by analysis of clones obd¹ by PCR amplification (42
 CC clones in total). The NANBH HC-J8 genome was found to contain a
 CC degenerate open reading frame encoding polypeptide precursors of 3033
 CC amino acid residues. See also AAR3358-9. (Updated on 25-MAR-2003 to
 CC correct PN field.) (Updated on 27-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 3033 AA:
 Query Match 85.6%; Score 13805; DB 2; Length 3033;
 Best Local Similarity 83.3%; Pred. No. 0;
 Matches 2525; Conservative 246; Mismatches 262; Indels 0; Gaps 0;

QY 301 VDDCNCSITPGTITGHRNAMDMMNNSPTATNITLAVNRVPEVIDITISGAHWGMFGIA 360
 DB 301 TOECNCSITGCHITGHRNAMDMMNNSPTATNITLAVNRVPEVIDITISGAHWGMFGIA 360
 QY 361 YFSMOGAAKAVVILLILLAGVDARHTHTVGSAAGQTGRLTSLFDMGPRCKQLQVNTNSW 420
 DB 361 YFSMOGAAKAVAILLLVAGVDASTYSTGOQAGRAAYGISLFTNGAQNHLITNNSW 420
 QY 421 HINRTALNCDLSHTGFIASLFYTHSFNNSGCPERRMSACRSIEAFRVMGALQVEDNYTN 480
 DB 421 HINRTALNCDLSLETGFIASLVYRRFNSSGCPERRLSGRLDDPRIGMTLETETNTYN 480
 QY 481 PEDMRPYCHNHPROCGVSAKTVCGPYCFFPSVVVYGTDRCLAPPTTGSENTDVL 540
 DB 481 DEDMRPYCHNHPRCGIVPATTGCPYVCFPSVVVYGTDRCLAPPTTGSENTDVL 540
 QY 541 LNSTRPPLGSMWFGCMNNSGVTKTGAPPCGTRADFNASTDLCPCTCFRRGPTTYLK 600
 DB 541 LNSTRPPRGAMFGKTMNGTFTKTKGAPPCIRKDYNSTIDLCPCTCFRRGPTTYLK 600
 QY 601 CGSGPMLTPRCLIDYPRLMHYPTVNTYTFKIRMYGVGVEHRLTAACNFTGRDNLSD 660
 DB 601 CGAGPMLTPRCLVDYPRLMHYPTVNTYTFKIRMYGVGVEHRLTAACNFTGRDNLSD 660
 QY 661 RDRQSPLHSTHTMALPCSYDLPALSTGLHLHQNIVDVQMTGSLPALTYIRW 720
 DB 661 RDRGQSPHLHSTHTMAVPCSPDLPALSTGLHLHQNIVDVQMTGSLPALTYIRW 720
 QY 721 EWNILLFLIADARVACILMMLILGOAEALAEKLVILHAASASONGFLYVIFVAVW 780
 DB 721 EWNILLFLIADARVACILMMLILGOAEALAEKLVILHAASASANGFLYVIFVAVW 780
 QY 781 YIKGRVPLATYSLTGLNFSLLILALPOAAVAYDASVHGOALVMTLFTLTPGX 840
 DB 781 YIKGRVPLATYSLTGLNFSLLILALPOAAVAYDASVHGOALVMTLFTLTPGX 840
 QY 841 TILSRFLMWLCYLLVLGEAMVDEMAPPMQVRGGRGIIWANAIFPGVVPDITTKLAVL 900
 DB 841 TILSRVWMLSTWLVAEKOIQWVPLEVRGGRGIIWANAIFPGVVPDITTKLAVL 900
 QY 901 GPAYILLKGLATVPFVBAHALRMCMTABRHLAGRVOMALTAGRTGTIYDHLTPM 960
 DB 901 GPAYILLKGLATVPFVBAHALRMCMTABRHLAGRVOMALTAGRTGTIYDHLTPM 960
 QY 961 SDMAASGLDLAAVAEPIIFSPMEKKVITWGAETACGDIIHLGVLVSARLREVLGPA 1020
 DB 961 STMAAQSRLDLAAVAEPIIFSPMEKKVITWGAETACGDIIHLGVLVSARLREVLGPA 1020
 QY 1021 GYTSKGMSLAPITTAQAOTRGLLGTIVYSMTGRDQAGETIOVLSVTQSLGTSISG 1080
 DB 1021 GYTSKGMMLAPITTAQAOTRGLLGTIVYSMTGRDQAGETIOVLSVTQSLGTSISG 1080
 QY 1081 VLMTYVHAGNKTILAGSRGPVTOMYSABGDLVGPSPPTGSLSPCGADVLYLRN 1140
 DB 1081 VLMTYVHAGNKTILAGSRGPVTOMYSABGDLVGPSPPTGSLSPCGADVLYLRN 1140
 QY 1141 ADVIPARRRGDRGALLSPRLSTLKSGSGGVLCPRHAGVGFRAAVCSRGVAKSIDFI 1200
 DB 1141 ADVIPARRRGDRGALLSPRLSTLKSGSGGVLCPRHAGVGFRAAVCSRGVAKSIDFI 1200
 QY 1201 PYETLIDYTRSPEDNSNPRAVPOTVOGYLHAATGSGKTKVVAAYAAOQYKLVNLP 1260
 DB 1201 PYESLIDYTRSPEDNSNPRAVPOTVOGYLHAATGSGKTKVVAAYAAOQYKLVNLP 1260
 QY 1261 SVAAATLGFAGYLSKAGHINPRTGRTVTTGAPRTTYGKFLADGGCAGAYDIIICD 1320
 DB 1261 SVAAATLGFAGYLSKAGHINPRTGRTVTTGAPRTTYGKFLADGGCAGAYDIIICD 1320
 QY 1321 ECHAVDSITLIGITVLDOAETAGVRLTVLATATPPGAVTTPHPIIEVALQGESEIPY 1380
 DB 1321 ECHAVDSITLIGITVLDOAETAGVRLTVLATATPPGAVTTPHPIIEVALQGESEIPY 1380
 QY 1381 GRAIPLSYIKGRHLIFCHSKKKCELLAALRGKGLNSVAYIRGLDVSIVIPQGDVVVA 1440

```

Db 1381 GAHIAPIAKGSHLIFCHSKKKCDLAAALRGTVANAVAYRGUDVSVIPIFGDVVVVA 1440
Qy 1441 TDAIMTGYGDDPSVDICNVAVTQVVDLSLDPFTTTTQIVPODANSRQRRTRGRL 1500
Db 1441 TDAIMTGYGDDPSVDICNVAVSQIVDFSLDPFTTTTQIVPODANSRQRRTRGRL 1500
Qy 1501 GYRYVSTGBRAGMDSVVLCECYDAGAAMELTPESETTVLRRAFNPGJLVCDHJE 1560
Db 1501 GYRYVSSGBRAGMDSVVPCECYDAGAAMELTPESETTVLRRAFNPGJLVCDHJE 1560
Qy 1561 FMEAVFTGLTHIDAHFLSOTKOSGENFAVLTAYQATVCARAKAPPSMDVMKCLTRLKP 1620
Db 1561 FMEAVFTGLTHINAHFLSQTKOGENFALITAYQATVCARAKAPPSMDVMKCLTRLKP 1620
Qy 1621 TLVGPFPLLYRLGUSTNEVTLTHPTVTKYIATCMQADLEWMTSVTLAGVLAANAACLA 1680
Db 1621 TLVGPFPLLYRLGAVNEVTLTHPTVTKYIATCMQADLEWMTSVTLAGVLAANAACLA 1680
Qy 1681 TGCVCIIIGRLHINORAVVAPDKEVLYEARFEMEBCASRAALIEGORIAEMLSKTIQGLL 1740
Db 1681 TGCISIIIGRLHNDRVVVPDKEILYEARFEMEBCASKALLIEGORMBMLSKTIQGLL 1740
Qy 1741 QOASKOADIOPVQASWPKVEQFMAKHMNFISGIOYLAGLSTLPGNPAVASMASFSA 1800
Db 1741 QOATRAQOGMOPALOSWPKLBOFMAKHMNFISGIOYLAGLSTLPGNPAVASMASFSA 1800
Qy 1801 LNSPLSTSTIIILNIGMULASQIAPPAATGTFVSGLVGAAGVSTIGLQKVLVDIILAGYG 1860
Db 1801 LNSPLSTSTIIILNIGMULASQIAPPAATGTFVSGLVGAAGVSTIGLQKVLVDIILAGYG 1860
Qy 1861 AGISGLVAFKIMSGKSPMEDVYNLLPGLISPGALVGVYICAILRRHVGPBGAVOMV 1920
Db 1861 AGISGLVAFKIMSGKSPMEDVYNLLPGLISPGALVGVYICAILRRHVGPBGAVOMV 1920
Qy 1921 NBLIAPASRGNHVAPTHYVTESDASQVTLQSLTITSLRLHNMITEDCPIPCGSGM 1980
Db 1921 NBLIAPASRGNHVAPTHYVTESDASQVTLQSLTITSLRLHNMITEDCPIPCGSGM 1980
Qy 1981 LQDIDMWCSTIITDFPKMILSSKILPMPGIPFISCKGKGVAGVWTRRPCANIS 2040
Db 1981 LQDIDMWCSTIITDFPKMILSSKILPMPGIPFISCKGKGVAGVWTRRPCANIS 2040
Qy 2041 GNVRLGSMRITGPKTCMNIWOGFPINCTEGOCVCKPAPNFVAILWRVAASRYAVTOH 2100
Db 2041 GNVRLGSMRITGPKTCMNIWOGFPINCTEGOCVCKPAPNFVAILWRVAASRYAVTOH 2100
Qy 2101 GSYHYITGLTTLNKLKVPQCLPSPFPFWSWDGVOIHRFAFPTPKPFEDFVSFCVGLNSFVY 2160
Db 2101 GSFYSYITGLTSDNLKVPQCVPAPEFBSWDGVOIHRFAFVPGFPRDEVTFTVGLNSFVY 2160
Qy 2161 GSQLPDPEPDDVILMSMLTDBSHITAETPAARLARSPSEASSASQLSAPSLRATCT 2220
Db 2161 GSQLPDPEPDDVILMSMLTDBSHITAETPAARLARSPSEASSASQLSAPSLRATCT 2220
Qy 2221 THGKAUVUMDANLPMGDDVTRIESGSKVVLVDSLDPWVERSDLEPSIPSYMLPKKR 2280
Db 2221 THGKAUVUMDANLPMGDDVTRIESGSKVVLVDSLDPWVERSDLEPSIPSYMLPKKR 2280
Qy 2281 FPPALPAMARPDYNPPLVESWKRPDYQAPTVAGCALPPRKTPTPPRRRRFTVGLSBDSI 2340
Db 2281 FPPALPAMARPDYNPPLVESWKRPDYQAPTVAGCALPPRKTPTPPRRRRFTVGLSBDSI 2340
Qy 2341 GDLAQOLAKSPGPPSGDSGLSTGAGAADSQTPPDELALSERGSISSMPLEGEIG 2400
Db 2341 GDLAQOLAKSPGPPSGDSGLSTGAGAADSQTPPDELALSERGSISSMPLEGEIG 2400
Qy 2401 DPLBEQOYEPPOPPOGVAAPSGDSGMSSTCSBEDSVVCCSMYSMTGALITPCSPPE 2460
Db 2401 DPLBEQOYEPPOPPOGVAAPSGDSGMSSTCSBEDSVVCCSMYSMTGALITPCSPPE 2460
Qy 2461 EKLPIINPLSNSLIRYANKVYCTTTKASLRAKKVTRDMQVLDLSYDVSVDKIKLAASRY 2520

```

```

Db 2461 EKLPIINPLSNSLIRYANKVYCTTTKASLRAKKVTRDMQVLDLSYDVSVDKIKLAASRY 2520
Qy 2521 TARIITMEACOLTTPHSAKSYGGAKEVRSLSGRAVNHISWMDLLEDBETPIPTI 2580
Db 2521 GARLITVEBACALTTPHSAKSYGGAKEVRSLSGRAVNHISWMDLLEDBETPIPTI 2580
Qy 2581 MAKNEVFCVDPPIKGGKARLIVYDPLGVRVEKMAIYDITOKLPOAVGASGYEQYSPA 2640
Db 2581 MAKNEVFCVDPPIKGGKARLIVYDPLGVRVEKMAIYDITOKLPOAVGASGYEQYSPA 2640
Qy 2641 QREVELLKAMAERKDDPMGFSYDTRCDSTVTERDIRTEBSIYRACSLPEBAHTAHSLTE 2700
Db 2641 QREVELLKAMAERKDDPMGFSYDTRCDSTVTERDIRTEBSIYRACSLPEBAHTAHSLTE 2700
Qy 2701 RLYVGPMPNSKGOCCGYRCSASGLVLTSMKNTITCYKALAACCAAGIIPTMVYCD 2760
Db 2701 RLYVGPMPNSKGOCCGYRCSISGFTTSMKNTITCYKALAACCAAGIIPTMVYCD 2760
Qy 2761 DLVVISSEQTEDEBNILRAFTAMTRYSAPPGDPRPEYDELITSCSNVSVALGPOG 2820
Db 2761 DLVVISSEQTEDEBNILRAFTAMTRYSAPPGDPRPEYDELITSCSNVSVALGPOG 2820
Qy 2821 RRRYVLTDPPTPIARAAMETVRHSPVNSWLGNIIOYAFIWARVYLMTHFSSIIAODT 2880
Db 2821 RRRYVLTDPPTPIARAAMETVRHSPVNSWLGNIIOYAFIWARVYLMTHFSSIIAODT 2880
Qy 2881 LDONINFEYMGAVYSPLDPAIIRLHGLDAFSLHTTTPHELTRVSAALTKGAPLR 2940
Db 2881 LDONINFEYMGAVYSPLDPAIIRLHGLDAFSLHTTTPHELTRVSAALTKGAPLR 2940
Qy 2941 AKMSRARAARASLISRGGRAAVCGRYLFMAVAKTKLKTPLPEARLDDLSMFTVAGAGG 3000
Db 2941 AKMSRARAARASLISRGGRAAVCGRYLFMAVAKTKLKTPLPEARLDDLSMFTVAGAGG 3000
Qy 3001 DIYHSVSRARPRLLFGLILFLVGVGLFLPAR 3033
Db 3001 DIYHSVSRARPRLLFGLILFLVGVGLFLPAR 3033

RESULT 9
AAB30730
ID AAB30730 standard; protein; 3015 AA.
XX
AC AAB30730;
XX
DT 02-APR-2001 (first entry)
XX
DE Amino acid sequence of chimeric Hepatitis C virus clone pH77CV-J65.
XX
KW HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine.
XX
OS Synthetic.
XX
OS Hepatitis C virus.
XX
PN MO200075338-A2.
XX
PD 14-DEC-2000.
XX
PF 02-JUN-2000; 2000WO-US015446.
XX
PR 04-JUN-1999; 99US-0137693P.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Yanagi M, Bukh J, Emerson SU, Purcell RH;
XX
DR WPI; 2001-061728/07.
XX
DR N-PSDB; AAC86645.
XX
PT Nucleic acid molecule encoding human hepatitis C virus of genotype 2a for
XX
PT developing vaccines, for diagnosis of hepatitis C virus and in screening
XX
PT assays for identification of antiviral agents.
XX

```

PS Disclosure; Page 103-115; 167pp; English.

XX AAB30730-33 are encoded by chimeric cDNA clones of infectious Hepatitis C virus (HCV). In each clone the C, E1 and E2 genes are derived from HCV strain HC-J6CH genotype 2a. The p7 protein was derived from HCV strain HC-J6CH or HCV strain PCV-H77C, and the NS genes were derived from HCV strain PCV-H77C. Such HCV sequences are capable of expressing the virus when transfected into cells. The HCV protein is useful for assaying candidate antiviral agents for activity against HCV. Antibodies specific for HCV polypeptide are useful in prevention and treatment of diseases caused by HCV in animals, in particular humans. The HCV polypeptides serve as immunogens in the development of vaccines for preventing HCV in mammals or as antigens in diagnostic assays for detecting the presence of HCV in biological samples. The HCV polynucleotide is also useful for identifying cell lines capable of supporting the replication of HCV in vitro and to produce attenuated viral strains via passage in vitro or in vivo

XX Sequence 3015 AA;

SQ

Query Match 80.9%; Score 13049; DB 4; Length 3015; Best Local Similarity 78.8%; Pred. No. 0; Matches 2395; Conservative 261; Mismatches 353; Indels 30; Gaps 4;

QY 1 MSTNPKQRTKNTKNTNRRPODVKFGGGQIVGVYLLPRRGPLGVATRTKTSERQPRG 60

DB 1 MSTNPKQRTKNTKNTNRRPODVKFGGGQIVGVYLLPRRGPLGVATRTKTSERQPRG 60

QY 61 RROPIPDORSSTKSMKPGYPMPLVGNBGLGNAQMLSPRGRPSGMPDPHRRRNVC 120

DB 61 RROPIPDORSSTKSMKPGYPMPLVGNBGLGNAQMLSPRGRPSGMPDPHRRRNVC 120

QY 121 KVIDTLTLCGFADLMGVIPVVGAPLGGVAAALAHGVRLSDGVNPAFGNLPGCSFSIFLLA 180

DB 121 KVIDTLTLCGFADLMGVIPVVGAPLGGVAAALAHGVRLSDGVNPAFGNLPGCSFSIFLLA 180

QY 121 KVIDTLTLCGFADLMGVIPVVGAPLGGVAAALAHGVRLSDGVNPAFGNLPGCSFSIFLLA 180

DB 121 KVIDTLTLCGFADLMGVIPVVGAPLGGVAAALAHGVRLSDGVNPAFGNLPGCSFSIFLLA 180

QY 181 LLSCTTPVSAAEVKNIISTGYMTNDCTNDSTIMOLAAVLHAPGCPVCKVGNASQCMW 240

DB 181 LLSCTTPVSAAEVKNIISTGYMTNDCTNDSTIMOLAAVLHAPGCPVCKVGNASQCMW 240

QY 181 LLSCTTPVSAAEVKNIISTGYMTNDCTNDSTIMOLAAVLHAPGCPVCKVGNASQCMW 240

DB 181 LLSCTTPVSAAEVKNIISTGYMTNDCTNDSTIMOLAAVLHAPGCPVCKVGNASQCMW 240

QY 241 PVSNNVAVORPGALTOGLRTHIDMVMSATLCSALVYGDLCGGVMLAQMFIYSPOHMF 300

DB 241 PVSNNVAVORPGALTOGLRTHIDMVMSATLCSALVYGDLCGGVMLAQMFIYSPOHMF 300

QY 241 PVSNNVAVORPGALTOGLRTHIDMVMSATLCSALVYGDLCGGVMLAQMFIYSPOHMF 300

DB 241 PVSNNVAVORPGALTOGLRTHIDMVMSATLCSALVYGDLCGGVMLAQMFIYSPOHMF 300

QY 301 VODNCISYPTGTTGHRMADMMNMSPTATMLAYMRVPEYIIDISGAHMGVFGFLA 360

DB 301 VODNCISYPTGTTGHRMADMMNMSPTATMLAYMRVPEYIIDISGAHMGVFGFLA 360

QY 301 VODNCISYPTGTTGHRMADMMNMSPTATMLAYMRVPEYIIDISGAHMGVFGFLA 360

DB 301 VODNCISYPTGTTGHRMADMMNMSPTATMLAYMRVPEYIIDISGAHMGVFGFLA 360

QY 361 YFSMOGMAKVVVILLAAAGVDARTHTVGGSAQTTGRLTSLFDMGRPKIQLVNTGSM 420

DB 361 YFSMOGMAKVVVILLAAAGVDARTHTVGGSAQTTGRLTSLFDMGRPKIQLVNTGSM 420

QY 361 YFSMOGMAKVVVILLAAAGVDARTHTVGGSAQTTGRLTSLFDMGRPKIQLVNTGSM 420

DB 361 YFSMOGMAKVVVILLAAAGVDARTHTVGGSAQTTGRLTSLFDMGRPKIQLVNTGSM 420

QY 421 HIRNTALNCNDSLHTGFIASLFYTHSNSSGCCPERMSACRSIAFRGKALDYEDNVN 480

DB 421 HIRNTALNCNDSLHTGFIASLFYTHSNSSGCCPERMSACRSIAFRGKALDYEDNVN 480

QY 421 HIRNTALNCNDSLHTGFIASLFYTHSNSSGCCPERMSACRSIAFRGKALDYEDNVN 480

DB 421 HIRNTALNCNDSLHTGFIASLFYTHSNSSGCCPERMSACRSIAFRGKALDYEDNVN 480

QY 481 PEDMRPYCMHYPPROCGVNSAKTVCGPVYCTFSPVNVGTTDLGAFTYTMGENETDVEL 540

DB 481 PEDMRPYCMHYPPROCGVNSAKTVCGPVYCTFSPVNVGTTDLGAFTYTMGENETDVEL 540

QY 481 PEDMRPYCMHYPPROCGVNSAKTVCGPVYCTFSPVNVGTTDLGAFTYTMGENETDVEL 540

DB 481 PEDMRPYCMHYPPROCGVNSAKTVCGPVYCTFSPVNVGTTDLGAFTYTMGENETDVEL 540

QY 541 LNSTRPPLGSMFGCTMMNSGTYTKCGAPPCRTAADFNASTDLCPTDCRKPDDTYLK 600

DB 541 LNSTRPPLGSMFGCTMMNSGTYTKCGAPPCRTAADFNASTDLCPTDCRKPDDTYLK 600

QY 541 LNSTRPPLGSMFGCTMMNSGTYTKCGAPPCRTAADFNASTDLCPTDCRKPDDTYLK 600

DB 541 LNSTRPPLGSMFGCTMMNSGTYTKCGAPPCRTAADFNASTDLCPTDCRKPDDTYLK 600

QY 601 CGSGPMLTPRCLIDYPRILMHYPTVYVYTI FKIRMYGVGVHRLTAACNTRDRCKLE 660

DB 601 CGSGPMLTPRCLIDYPRILMHYPTVYVYTI FKIRMYGVGVHRLTAACNTRDRCKLE 660

QY 601 CGSGPMLTPRCLIDYPRILMHYPTVYVYTI FKIRMYGVGVHRLTAACNTRDRCKLE 660

DB 601 CGSGPMLTPRCLIDYPRILMHYPTVYVYTI FKIRMYGVGVHRLTAACNTRDRCKLE 660

QY 661 RDRSOLSPLLHSTTEMAILPCSYSDLPALSTGLHLHQNIVDVQFMYGSLPALTKYIVR 720

DB 661 RDRSOLSPLLHSTTEMAILPCSYSDLPALSTGLHLHQNIVDVQFMYGSLPALTKYIVR 720

QY 661 RDRSOLSPLLHSTTEMAILPCSYSDLPALSTGLHLHQNIVDVQFMYGSLPALTKYIVR 720

DB 661 RDRSOLSPLLHSTTEMAILPCSYSDLPALSTGLHLHQNIVDVQFMYGSLPALTKYIVR 720

QY 721 EWWILLFLLLADARVACIMMLILLGOAEBALBKLVTILHAASAACNGFLYFVIFVAVM 780

DB 721 EWWILLFLLLADARVACIMMLILLGOAEBALBKLVTILHAASAACNGFLYFVIFVAVM 780

QY 781 YIKGVRVPLATYSLGLMSFSLILLALPQOAAVYASVHGQIGALLWITLFTLPQK 840

DB 781 YIKGVRVPLATYSLGLMSFSLILLALPQOAAVYASVHGQIGALLWITLFTLPQK 840

QY 841 TLLSFFLMLCYLLTLGEMVEMAPPMQVGRGDGIIWAVAFYPGVFDITKMLAVL 900

DB 841 TLLSFFLMLCYLLTLGEMVEMAPPMQVGRGDGIIWAVAFYPGVFDITKMLAVL 900

QY 841 TLLSFFLMLCYLLTLGEMVEMAPPMQVGRGDGIIWAVAFYPGVFDITKMLAVL 900

DB 841 TLLSFFLMLCYLLTLGEMVEMAPPMQVGRGDGIIWAVAFYPGVFDITKMLAVL 900

QY 901 GRAYLLKALTRVFPVRAHALLMCTMARHLAGRYQMALALGRMTGTITYHILPM 960

DB 901 GRAYLLKALTRVFPVRAHALLMCTMARHLAGRYQMALALGRMTGTITYHILPM 960

QY 901 GRAYLLKALTRVFPVRAHALLMCTMARHLAGRYQMALALGRMTGTITYHILPM 960

DB 901 GRAYLLKALTRVFPVRAHALLMCTMARHLAGRYQMALALGRMTGTITYHILPM 960

QY 961 SDMAASGRDLVAVEPIIFSPMEKKVIVMGETAACGDIHLGFPVSARLREVLGPAD 1020

DB 961 SDMAASGRDLVAVEPIIFSPMEKKVIVMGETAACGDIHLGFPVSARLREVLGPAD 1020

QY 961 SDMAASGRDLVAVEPIIFSPMEKKVIVMGETAACGDIHLGFPVSARLREVLGPAD 1020

DB 961 SDMAASGRDLVAVEPIIFSPMEKKVIVMGETAACGDIHLGFPVSARLREVLGPAD 1020

QY 1021 GYTSKQMSLLAPITAYAOOTRGLLGTIVVSMTRDTEQAGEIQLVSTVTSFLOTISIG 1080

DB 1021 GYTSKQMSLLAPITAYAOOTRGLLGTIVVSMTRDTEQAGEIQLVSTVTSFLOTISIG 1080

QY 1021 GYTSKQMSLLAPITAYAOOTRGLLGTIVVSMTRDTEQAGEIQLVSTVTSFLOTISIG 1080

DB 1021 GYTSKQMSLLAPITAYAOOTRGLLGTIVVSMTRDTEQAGEIQLVSTVTSFLOTISIG 1080

QY 1081 VLTNTYHGAQNTLAGSGRPVTQMTSAEGDLVGNPSPPTKSLPCTCGAVDLVLRN 1140

DB 1081 VLTNTYHGAQNTLAGSGRPVTQMTSAEGDLVGNPSPPTKSLPCTCGAVDLVLRN 1140

QY 1081 VLTNTYHGAQNTLAGSGRPVTQMTSAEGDLVGNPSPPTKSLPCTCGAVDLVLRN 1140

DB 1081 VLTNTYHGAQNTLAGSGRPVTQMTSAEGDLVGNPSPPTKSLPCTCGAVDLVLRN 1140

QY 1141 ADVIPARRGDKGALSPRPSTLKGSSGGVLCPRGNAVGFPAACSRVNASIFPT 1200

DB 1141 ADVIPARRGDKGALSPRPSTLKGSSGGVLCPRGNAVGFPAACSRVNASIFPT 1200

QY 1141 ADVIPARRGDKGALSPRPSTLKGSSGGVLCPRGNAVGFPAACSRVNASIFPT 1200

DB 1141 ADVIPARRGDKGALSPRPSTLKGSSGGVLCPRGNAVGFPAACSRVNASIFPT 1200

QY 1201 PYETLIDVTRSPFPDSNSTPPAVPQTYGYLHAPTSQSKSTKVPAAYAAQGYKLVLP 1260

DB 1201 PYETLIDVTRSPFPDSNSTPPAVPQTYGYLHAPTSQSKSTKVPAAYAAQGYKLVLP 1260

QY 1201 PYETLIDVTRSPFPDSNSTPPAVPQTYGYLHAPTSQSKSTKVPAAYAAQGYKLVLP 1260

DB 1201 PYETLIDVTRSPFPDSNSTPPAVPQTYGYLHAPTSQSKSTKVPAAYAAQGYKLVLP 1260

QY 1261 SYAATLGRAYLSKAGINPNRTGVRVTTGAPITTYGTYGKFLADGCGAGAYDIIICD 1320

DB 1261 SYAATLGRAYLSKAGINPNRTGVRVTTGAPITTYGTYGKFLADGCGAGAYDIIICD 1320

QY 1261 SYAATLGRAYLSKAGINPNRTGVRVTTGAPITTYGTYGKFLADGCGAGAYDIIICD 1320

DB 1261 SYAATLGRAYLSKAGINPNRTGVRVTTGAPITTYGTYGKFLADGCGAGAYDIIICD 1320

QY 1321 ECHAVDSTIIIGIGVLDQAEFAGVRLTVLATATPPGSLTPHPNIBEVALQEGEIPY 1380

DB 1321 ECHAVDSTIIIGIGVLDQAEFAGVRLTVLATATPPGSLTPHPNIBEVALQEGEIPY 1380

QY 1321 ECHAVDSTIIIGIGVLDQAEFAGVRLTVLATATPPGSLTPHPNIBEVALQEGEIPY 1380

DB 1321 ECHAVDSTIIIGIGVLDQAEFAGVRLTVLATATPPGSLTPHPNIBEVALQEGEIPY 1380

QY 1381 GRAIPLSYIKGGRHLIFCHSKKKCDLAAALGKGLNSVAYYRGLDVSVIPTQSDVVVA 1440

DB 1381 GRAIPLSYIKGGRHLIFCHSKKKCDLAAALGKGLNSVAYYRGLDVSVIPTQSDVVVA 1440

QY 1381 GRAIPLSYIKGGRHLIFCHSKKKCDLAAALGKGLNSVAYYRGLDVSVIPTQSDVVVA 1440

DB 1381 GRAIPLSYIKGGRHLIFCHSKKKCDLAAALGKGLNSVAYYRGLDVSVIPTQSDVVVA 1440

QY 1441 TDALMTGYTGDPSYIDCNVAVTOVVDPSLDPTFTITTIQIVPDAVSQRGRGTGRGL 1500

DB 1441 TDALMTGYTGDPSYIDCNVAVTOVVDPSLDPTFTITTIQIVPDAVSQRGRGTGRGL 1500

QY 1441 TDALMTGYTGDPSYIDCNVAVTOVVDPSLDPTFTITTIQIVPDAVSQRGRGTGRGL 1500

DB 1441 TDALMTGYTGDPSYIDCNVAVTOVVDPSLDPTFTITTIQIVPDAVSQRGRGTGRGL 1500

QY 1501 GIYRVYSTERASGMFDSVVLCECYDAGAAWELTPSETTVLRAYFPTPGLPVQDHL 1560

DB 1501 GIYRVYSTERASGMFDSVVLCECYDAGAAWELTPSETTVLRAYFPTPGLPVQDHL 1560

QY 1501 GIYRVYSTERASGMFDSVVLCECYDAGAAWELTPSETTVLRAYFPTPGLPVQDHL 1560

DB 1501 GIYRVYSTERASGMFDSVVLCECYDAGAAWELTPSETTVLRAYFPTPGLPVQDHL 1560

QY 1561 FWEAVFTGLTHIDAFLSQTKQSGENFAYLTAAYQATVCARAKAPPSMDVMKCLTRLKP 1620

DB 1561 FWEAVFTGLTHIDAFLSQTKQSGENFAYLTAAYQATVCARAKAPPSMDVMKCLTRLKP 1620

QY 1561 FWEAVFTGLTHIDAFLSQTKQSGENFAYLTAAYQATVCARAKAPPSMDVMKCLTRLKP 1620

DB 1561 FWEAVFTGLTHIDAFLSQTKQSGENFAYLTAAYQATVCARAKAPPSMDVMKCLTRLKP 1620

QY 1621 TLVGPPLLYRIGSTYNTBTLTHPTKYIATCMODLEMTSTTWLAGVLAAYACLA 1680

DB 1621 TLVGPPLLYRIGSTYNTBTLTHPTKYIATCMODLEMTSTTWLAGVLAAYACLA 1680

QY 1621 TLVGPPLLYRIGSTYNTBTLTHPTKYIATCMODLEMTSTTWLAGVLAAYACLA 1680

DB 1621 TLVGPPLLYRIGSTYNTBTLTHPTKYIATCMODLEMTSTTWLAGVLAAYACLA 1680

QY 1681 TGCVCIGRLHINOAVVAPKEVLYEAFDEMECASRALILEGQRIAEMLSKIQGLL 1740

DB 1681 TGCVCIGRLHINOAVVAPKEVLYEAFDEMECASRALILEGQRIAEMLSKIQGLL 1740

QY 1681 TGCVCIGRLHINOAVVAPKEVLYEAFDEMECASRALILEGQRIAEMLSKIQGLL 1740

DB 1681 TGCVCIGRLHINOAVVAPKEVLYEAFDEMECASRALILEGQRIAEMLSKIQGLL 1740

QY 1741 QOASRQADIOPTVOASWPKQEFWAKIMNFIISGIIYLAGISTLPGNPAVASMAFSA 1800

DB 1741 QOASRQADIOPTVOASWPKQEFWAKIMNFIISGIIYLAGISTLPGNPAVASMAFSA 1800

QY 1741 QOASRQADIOPTVOASWPKQEFWAKIMNFIISGIIYLAGISTLPGNPAVASMAFSA 1800

DB 1741 QOASRQADIOPTVOASWPKQEFWAKIMNFIISGIIYLAGISTLPGNPAVASMAFSA 1800

QY 1801 LTPSLSTSTIIILNLGMLASQIAPPAGATGFVSVGLGAAVGSIGLKVYDLIAGV 1860

DB 1801 LTPSLSTSTIIILNLGMLASQIAPPAGATGFVSVGLGAAVGSIGLKVYDLIAGV 1860

QY 1801 LTPSLSTSTIIILNLGMLASQIAPPAGATGFVSVGLGAAVGSIGLKVYDLIAGV 1860

DB 1801 LTPSLSTSTIIILNLGMLASQIAPPAGATGFVSVGLGAAVGSIGLKVYDLIAGV 1860

```

Db      1801 VTSPLTTGQTLLFNILGWNVAQAALAPGATATFVAGLAGAAIGSVGLKVLVDILAGYG 1860
Qy      1861 AGISGALVAFKIMSGEKSPMEDVYVNLPGILSPGALVGVICAILIRRHVGEGEAVQNM 1920
Db      1861 AAVAGALVAFKIMSGEVPSTEDLVNLPAILSPGALVGVICAILIRRHVGEGEAVQNM 1920
Qy      1921 NRIILAFASRGHVAPTHVYTESDASQVQLGSLITITSLIRLHMWITEDCPIPCQSGM 1980
Db      1921 NRIILAFASRGHVAPTHVYTESDASQVQLGSLITITSLIRLHMWITEDCPIPCQSGM 1980
Qy      1981 LRDVMDVCTIITLDFKNMLTSKLPFPMGSLPVYSCQKYGKVGWAGTGIMTTRCPGANIS 2040
Db      1981 LRDVMDVCTIITLDFKNMLTSKLPFPMGSLPVYSCQKYGKVGWAGTGIMTTRCPGANIS 2040
Qy      2041 GNVRLGSMRITGPKTCMNIMOGFPPINCTEGOCVCKPAPNPFVAIAWRVAASEYAEVTOH 2100
Db      2041 GNVRLGSMRITGPKTCMNIMOGFPPINCTEGOCVCKPAPNPFVAIAWRVAASEYAEVTOH 2100
Qy      2101 GSVHYITGLTTNLLKVPQCLPSPEFSSWVGVOIHRFAPTPKPFDEVSFCVGLNSFVY 2160
Db      2101 GSVHYITGLTTNLLKVPQCLPSPEFSSWVGVOIHRFAPTPKPFDEVSFCVGLNSFVY 2160
Qy      2161 GSQLPCEPDPDVLVMSMLTDPSHITAEIAPARLARAGSPPESSASSASQLSAPSLRATCT 2220
Db      2161 GSQLPCEPDPDVLVMSMLTDPSHITAEIAPARLARAGSPPESSASSASQLSAPSLRATCT 2220
Qy      2221 THGKAYDVDMVANLF---MGDVTIRISSGKVVVLDLDPVVERSRDLSEIPSEVNL 2276
Db      2221 THGKAYDVDMVANLF---MGDVTIRISSGKVVVLDLDPVVERSRDLSEIPSEVNL 2276
Qy      2276 THGKAYDVDMVANLF---MGDVTIRISSGKVVVLDLDPVVERSRDLSEIPSEVNL 2276
Db      2276 THGKAYDVDMVANLF---MGDVTIRISSGKVVVLDLDPVVERSRDLSEIPSEVNL 2276
Qy      2277 PKKRPFPALPAMARPDYNPPLVESWKRPDYOPATVAGCALPPPKRPTPPRRRRRVGLS 2336
Db      2277 PKKRPFPALPAMARPDYNPPLVESWKRPDYOPATVAGCALPPPKRPTPPRRRRRVGLS 2336
Qy      2336 PKKRPFPALPAMARPDYNPPLVESWKRPDYOPATVAGCALPPPKRPTPPRRRRRVGLS 2336
Db      2336 PKKRPFPALPAMARPDYNPPLVESWKRPDYOPATVAGCALPPPKRPTPPRRRRRVGLS 2336
Qy      2337 EDSIGDALQQLAIKSFQOPPSGDSGLSTGAGAAGSGQTPDELALSETGISMPLE 2396
Db      2337 EDSIGDALQQLAIKSFQOPPSGDSGLSTGAGAAGSGQTPDELALSETGISMPLE 2396
Qy      2396 EDSIGDALQQLAIKSFQOPPSGDSGLSTGAGAAGSGQTPDELALSETGISMPLE 2396
Db      2396 EDSIGDALQQLAIKSFQOPPSGDSGLSTGAGAAGSGQTPDELALSETGISMPLE 2396
Qy      2397 GBLGPDLEPEOVEPPOPPOGGVAAAGSGSGSSTGSE--EDDSVYCCSMYSMTALTT 2454
Db      2397 GBLGPDLEPEOVEPPOPPOGGVAAAGSGSGSSTGSE--EDDSVYCCSMYSMTALTT 2454
Qy      2454 GBLGPDLEPEOVEPPOPPOGGVAAAGSGSGSSTGSE--EDDSVYCCSMYSMTALTT 2454
Db      2454 GBLGPDLEPEOVEPPOPPOGGVAAAGSGSGSSTGSE--EDDSVYCCSMYSMTALTT 2454
Qy      2496 PCABEQKLPINALNSLIRHNLVSTTSRSRCQOKVYTRDLQVLSHODVLKEXY 2496
Db      2496 PCABEQKLPINALNSLIRHNLVSTTSRSRCQOKVYTRDLQVLSHODVLKEXY 2496
Qy      2515 LAASKVTAALLTMEACQLTPPHSAASKYFGAKVEYRSLSGAVNHIXSVMDLLEDSET 2574
Db      2515 LAASKVTAALLTMEACQLTPPHSAASKYFGAKVEYRSLSGAVNHIXSVMDLLEDSET 2574
Qy      2574 LAASKVTAALLTMEACQLTPPHSAASKYFGAKVEYRSLSGAVNHIXSVMDLLEDSET 2574
Db      2574 LAASKVTAALLTMEACQLTPPHSAASKYFGAKVEYRSLSGAVNHIXSVMDLLEDSET 2574
Qy      2575 PIPTTMAKNEVFCVDPPTGKKAARLIYYPDLGVRCERKMAIYITQKLPQAVMGASYG 2634
Db      2575 PIPTTMAKNEVFCVDPPTGKKAARLIYYPDLGVRCERKMAIYITQKLPQAVMGASYG 2634
Qy      2634 PIPTTMAKNEVFCVDPPTGKKAARLIYYPDLGVRCERKMAIYITQKLPQAVMGASYG 2634
Db      2634 PIPTTMAKNEVFCVDPPTGKKAARLIYYPDLGVRCERKMAIYITQKLPQAVMGASYG 2634
Qy      2663 PYSYPAORVEFLIKAAAEKKDPMGFSYDRCPDSTYTERDIRTESIYACSLPEBAHNA 2694
Db      2663 PYSYPAORVEFLIKAAAEKKDPMGFSYDRCPDSTYTERDIRTESIYACSLPEBAHNA 2694
Qy      2694 PYSYPAORVEFLIKAAAEKKDPMGFSYDRCPDSTYTERDIRTESIYACSLPEBAHNA 2694
Db      2694 PYSYPAORVEFLIKAAAEKKDPMGFSYDRCPDSTYTERDIRTESIYACSLPEBAHNA 2694
Qy      2695 IHSLETERLVYVGMFENSKQTCGRRCRASGVLTITSMGNTITCYVVALAACRAAGIATP 2754
Db      2695 IHSLETERLVYVGMFENSKQTCGRRCRASGVLTITSMGNTITCYVVALAACRAAGIATP 2754
Qy      2754 IHSLETERLVYVGMFENSKQTCGRRCRASGVLTITSMGNTITCYVVALAACRAAGIATP 2754
Db      2754 IHSLETERLVYVGMFENSKQTCGRRCRASGVLTITSMGNTITCYVVALAACRAAGIATP 2754
Qy      2677 IKSLETERLVYVGGPPLTNSRGENCGRRCRASGVLTITCGNTLTCYIKARAACRAAGIDCT 2736
Db      2677 IKSLETERLVYVGGPPLTNSRGENCGRRCRASGVLTITCGNTLTCYIKARAACRAAGIDCT 2736
Qy      2736 IKSLETERLVYVGGPPLTNSRGENCGRRCRASGVLTITCGNTLTCYIKARAACRAAGIDCT 2736
Db      2736 IKSLETERLVYVGGPPLTNSRGENCGRRCRASGVLTITCGNTLTCYIKARAACRAAGIDCT 2736
Qy      2755 MIVCGDDLVVISESQTEDERLRAFTEAMTRYKAPRQDPREPEVDELITSCSSNVSV 2814
Db      2755 MIVCGDDLVVISESQTEDERLRAFTEAMTRYKAPRQDPREPEVDELITSCSSNVSV 2814
Qy      2796 MIVCGDDLVVISESQTEDERLRAFTEAMTRYKAPRQDPREPEVDELITSCSSNVSV 2796
Db      2796 MIVCGDDLVVISESQTEDERLRAFTEAMTRYKAPRQDPREPEVDELITSCSSNVSV 2796
Qy      2815 ALGPGORRKYVLTTRDPTPIARAAMETVHSFVNSMIGNIOYAPITAMVMTATTFPFI 2874
Db      2815 ALGPGORRKYVLTTRDPTPIARAAMETVHSFVNSMIGNIOYAPITAMVMTATTFPFI 2874
Qy      2874 ALGPGORRKYVLTTRDPTPIARAAMETVHSFVNSMIGNIOYAPITAMVMTATTFPFI 2874
Db      2874 ALGPGORRKYVLTTRDPTPIARAAMETVHSFVNSMIGNIOYAPITAMVMTATTFPFI 2874
Qy      2797 AADGAKRKYVLTTRDPTPIARAAMETVHSFVNSMIGNIOYAPITAMVMTATTFPFI 2856
Db      2797 AADGAKRKYVLTTRDPTPIARAAMETVHSFVNSMIGNIOYAPITAMVMTATTFPFI 2856
Qy      2856 AADGAKRKYVLTTRDPTPIARAAMETVHSFVNSMIGNIOYAPITAMVMTATTFPFI 2856
Db      2856 AADGAKRKYVLTTRDPTPIARAAMETVHSFVNSMIGNIOYAPITAMVMTATTFPFI 2856
Qy      2875 IMAQDTLDONLNFEMYGAVYSVPLDPAIIEIRHGLDAFSLHTTYPHLETAVASARLU 2934
Db      2875 IMAQDTLDONLNFEMYGAVYSVPLDPAIIEIRHGLDAFSLHTTYPHLETAVASARLU 2934
Qy      2934 IMAQDTLDONLNFEMYGAVYSVPLDPAIIEIRHGLDAFSLHTTYPHLETAVASARLU 2934
Db      2934 IMAQDTLDONLNFEMYGAVYSVPLDPAIIEIRHGLDAFSLHTTYPHLETAVASARLU 2934
Qy      2857 LIARQDLBOALNCEIYGACYSIEPDLPIIQIRHGLDAFSLHTSYSGEINRVAACLRKL 2916
Db      2857 LIARQDLBOALNCEIYGACYSIEPDLPIIQIRHGLDAFSLHTSYSGEINRVAACLRKL 2916

```

```

Qy      2935 GAPPLRAMKRAVARASLISRGRAAVCGRYLFNMAVTKLKLPLPEARLLDISWPT 2994
Db      2917 GVPLPLAMRHRRARRARRALLISRGRAAIGSKLPLFMAVTKLKLPLIAAGRLDISWPT 2976
Qy      2995 VGAGGGDIYHSVSRARPLLFGLLLLFVGVLPFLPAR 3033
Db      2977 AGYSGGDIYHSVSHARPWFVFCLLLAAGVGITLLPNR 3015

RESULT 10
AAB30732
ID AAB30732 standard; protein; 3015 AA.
XX
AC AAB30732;
XX
DT 02-APR-2001 (first entry)
XX
DE Amino acid sequence of chimeric Hepatitis C virus clone J6S.
XX
KW HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine.
XX
OS Synthetic.
XX
OS Hepatitis C virus.
XX
PN MO200075338-A2.
XX
PD 14-DEC-2000.
XX
PF 02-JUN-2000; 2000WO-US015446.
XX
PR 04-JUN-1999; 99US-0137693P.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Yanagi M, Bukh J, Emerson SU, Purcell RH;
XX
DR WPI: 2001-061728/07.
XX
DR N-PSDB; AAC86647.
XX
PT Nucleic acid molecule encoding human hepatitis C virus of genotype 2a for
PT developing vaccines, for diagnosis of hepatitis C virus and in screening
PT assays for identification of antiviral agents.
XX
PS Disclosure; Page 134-146; 167pb; English.
XX
CC AAB30730-33 are encoded by chimeric cDNA clones of infectious Hepatitis C
CC virus (HCV). In each clone the C, E1 and E2 genes are derived from HCV
CC strain HC-J6CH genotype 2a. The p7 protein was derived from HCV strain HC
CC -J6CH or HCV strain PCV-H77C, and the NS genes were derived from HCV
CC strain PCV-H77C. Such HCV sequences are capable of expressing the virus
CC when transfectected into cells. The HCV protein is useful for assaying
CC candidate antiviral agents for activity against HCV. Antibodies specific
CC for HCV polypeptide are useful in prevention and treatment of diseases
CC caused by HCV in animals, in particular humans. The HCV polypeptides
CC serve as immunogens in the development of vaccines for preventing HCV in
CC mammals or as antigens in diagnostic assays for detecting the presence of
CC HCV in biological samples. The HCV polynucleotide is also useful for
CC identifying cell lines capable of supporting the replication of HCV in
CC vitro and to produce attenuated viral strains via passage in vitro or in
XX vivo
XX
SQ Sequence 3015 AA;
XX
Query Match 80.9%; Score 13049; DB 4; Length 3015;
Best Local Similarity 78.8%; Pred. No. 0;
Matches 2395; Conservative 261; Mismatches 353; Indels 30; Gaps 4;
Qy      1 MSTNRPQKTRNNRRRRPQDVKFPGGQIYGVGVLLPFRGRLVARRAKTSERSQPG 60
Db      1 MSTNRPQKTRNNRRRRPQDVKFPGGQIYGVGVLLPFRGRLVARRAKTSERSQPG 60
Qy      61 RQPIPKDRSRSTGKMGKRGYPWPLYGNEGGLQMAQMLSPRGSRPSWGPNDPRHRSRNVG 120

```

Db 61 RKOP1FKDRRSTGKSGKGPWPL1GNGBLGAGWMLSPRSGRPBGMGNDRHRSRNVG 120
 Qy 121 KYIDLTCGFADLMGVIPIVVGAPLGGVARALAGVVLSDGVNFAFAGN1PGGSFS1FLIA 180
 Db 121 KYIDLTCGFADLMGVIPIVVGAPLGGVARALAGVVLSDGVNFAFAGN1PGGSFS1FLIA 180
 Qy 181 LLSCLITTPVSAAEVKNIS1G1GVWVNDCTNDSITWQLOAAVLHVP6GCPCEKYGNASQCVI 240
 Db 181 LLSCLITTPVSAAEVKNIS1G1GVWVNDCTNDSITWQLOAAVLHVP6GCPCEKYGNASQCVI 240
 Qy 241 PVSBNVAVORPGAL1TGLKTHIDMVVMSATLCSALVYGDLCGGVMLAAMF1VSPQHMF 300
 Db 241 PVSBNVAVORPGAL1TGLKTHIDMVVMSATLCSALVYGDLCGGVMLAAMF1VSPQHMF 300
 Qy 301 VODNCISIYPTT1TGRHAMDMMNMSPTATMLAYAMR1PVIYDIIISGAHGVWFGIA 360
 Db 301 VODNCISIYPTT1TGRHAMDMMNMSPTATMLAYAMR1PVIYDIIISGAHGVWFGIA 360
 Qy 361 YESMOGAMAKVVILL1LAAGVADARTHTVGSSAAQTGR1LSLFDMPGPROK1Q1VNTNGSM 420
 Db 361 YESMOGAMAKVVILL1LAAGVADARTHTVGSSAAQTGR1LSLFDMPGPROK1Q1VNTNGSM 420
 Qy 421 H1NRTALNCNDS1HTGF1ASLFTTSHFNSSGCPBRMSACRS1EAF1PVGMA1QYEDNVTN 480
 Db 421 H1NRTALNCNDS1HTGF1ASLFTTSHFNSSGCPBRMSACRS1EAF1PVGMA1QYEDNVTN 480
 Qy 481 PEDMR1YCMHYP1ROCGVSAKTVGAPV1CPTPS1PVVGTTRDLG1PYTWMGBNEDVFL 540
 Db 481 PEDMR1YCMHYP1ROCGVSAKTVGAPV1CPTPS1PVVGTTRDLG1PYTWMGBNEDVFL 540
 Qy 541 LNSTREPLGSWFC2TMMNNSGYTKTCGAPPCRTADFNASTDL1CPTDCEFRKHDP1TYLK 600
 Db 541 LNSTREPLGSWFC2TMMNNSGYTKTCGAPPCRTADFNASTDL1CPTDCEFRKHDP1TYLK 600
 Qy 601 CGSGPMLTRCL1DYRYRLMHP1CTVNYT1FK1RMVGVGEHRL1AACNFTGRDCNLED 660
 Db 601 CGSGPMLTRCL1DYRYRLMHP1CTVNYT1FK1RMVGVGEHRL1AACNFTGRDCNLED 660
 Qy 661 RDRSOLSP1LHSTTEMA1LPCSYSD1PALSTGL1H1HONI1VDQFMYGSPALTKX1YVM 720
 Db 661 RDRSOLSP1LHSTTEMA1LPCSYSD1PALSTGL1H1HONI1VDQFMYGSPALTKX1YVM 720
 Qy 721 EWW1LLFLL1ADARVAC1LML1LLGOAEAL1EKL1YLHAAS1AAS1CNGFLYFV1FEV1AAM 780
 Db 721 EWW1LLFLL1ADARVAC1LML1LLGOAEAL1EKL1YLHAAS1AAS1CNGFLYFV1FEV1AAM 780
 Qy 781 Y1KGRVPLATYS1LGLMSFSL1LLALPOOAYAYDASVHGQ1GALL1WMT1LFTLT1PYK 840
 Db 781 Y1KGRVPLATYS1LGLMSFSL1LLALPOOAYAYDASVHGQ1GALL1WMT1LFTLT1PYK 840
 Qy 841 T1LSRFL1M1CYLL1LGEAMVQEM1APPMOYRGGRCG11MAVA1FYPGVVFD1TKML1AUL 900
 Db 841 T1LSRFL1M1CYLL1LGEAMVQEM1APPMOYRGGRCG11MAVA1FYPGVVFD1TKML1AUL 900
 Qy 901 GBAV1LLKGA1LTRV1FVRAH1AL1RMCTMAH1LAGRYVQMA1LALGRMTGY1YDHL1TBM 960
 Db 901 GBAV1LLKGA1LTRV1FVRAH1AL1RMCTMAH1LAGRYVQMA1LALGRMTGY1YDHL1TBM 960
 Qy 961 SDMA1SG1DL1A1V1EPI1FSPMEKKV1YWGAE1TAACG1LHGL1PV1SAR1LGR1V1L1GPAD 1020
 Db 961 SDMA1SG1DL1A1V1EPI1FSPMEKKV1YWGAE1TAACG1LHGL1PV1SAR1LGR1V1L1GPAD 1020
 Qy 1021 GYTSKMS1L1A1P1TA1A1O1TRG1L1G1T1V1SM1GR1DTE1BQ1AGE1Q1V1STV1Q1S1G1S1G 1080
 Db 1021 GYTSKMS1L1A1P1TA1A1O1TRG1L1G1T1V1SM1GR1DTE1BQ1AGE1Q1V1STV1Q1S1G1S1G 1080
 Qy 1081 V1MTV1Y1GA1GNK1T1AG1R1P1VT1Q1MS1A1BG1DL1V1G1MP1SP1G1T1K1L1E1B1CT1CG1A1DL1V1L1T1RN 1140
 Db 1081 V1MTV1Y1GA1GNK1T1AG1R1P1VT1Q1MS1A1BG1DL1V1G1MP1SP1G1T1K1L1E1B1CT1CG1A1DL1V1L1T1RN 1140
 Qy 1141 ADV1PARR1GDX1GALL1SP1PL1ST1LKG1SG1G1PV1L1C1PR1GA1V1G1FRA1V1C1SR1GV1A1S1D1PT 1200
 Db 1141 ADV1PARR1GDX1GALL1SP1PL1ST1LKG1SG1G1PV1L1C1PR1GA1V1G1FRA1V1C1SR1GV1A1S1D1PT 1200

Db 1141 ADV1PARR1GDX1GALL1SP1PL1ST1LKG1SG1G1PV1L1C1PR1GA1V1G1FRA1V1C1SR1GV1A1S1D1PT 1200
 Qy 1201 PVE1TDL1Y1R1SP1T1F1SD1N1ST1PP1AV1POT1YOV1Y1HA1PT1G1SG1K1V1V1A1A1O1GY1K1VL1V1NP 1260
 Db 1201 PVE1TDL1Y1R1SP1T1F1SD1N1ST1PP1AV1POT1YOV1Y1HA1PT1G1SG1K1V1V1A1A1O1GY1K1VL1V1NP 1260
 Qy 1261 SVA1AT1FG1AY1L1SKA1GIN1PNI1RTG1V1T1TG1AP1T1Y1ST1YG1K1FL1AD1GG1C1AGA1YD1I1ICD 1320
 Db 1261 SVA1AT1FG1AY1L1SKA1GIN1PNI1RTG1V1T1TG1AP1T1Y1ST1YG1K1FL1AD1GG1C1AGA1YD1I1ICD 1320
 Qy 1321 ECH1AVD1ST1I1G1GV1L1DQ1A1E1AG1R1VL1V1L1AT1P1RG1SV1T1PH1PI1E1VAL1G1B1E1PI1RY 1380
 Db 1321 ECH1AVD1ST1I1G1GV1L1DQ1A1E1AG1R1VL1V1L1AT1P1RG1SV1T1PH1PI1E1VAL1G1B1E1PI1RY 1380
 Qy 1381 GRA1PL1SY1K1G1GR1L1F1CH1SK1K1C1B1E1LA1K1V1A1G1M1A1V1A1Y1RG1D1V1I1P1SG1DV1V1S 1440
 Db 1381 GRA1PL1SY1K1G1GR1L1F1CH1SK1K1C1B1E1LA1K1V1A1G1M1A1V1A1Y1RG1D1V1I1P1SG1DV1V1S 1440
 Qy 1441 TDALMTG1Y1G1D1P1SD1V1D1CN1VAV1TQ1V1D1F1S1D1P1T1T1Q1V1P1D1AV1S1S1Q1RR1G1T1GR1L 1500
 Db 1441 TDALMTG1Y1G1D1P1SD1V1D1CN1VAV1TQ1V1D1F1S1D1P1T1T1Q1V1P1D1AV1S1S1Q1RR1G1T1GR1L 1500
 Qy 1501 G1Y1R1V1S1GER1AS1G1M1F1D1SV1L1CE1CYD1AC1AM1E1L1P1A1T1Y1V1L1A1Y1M1T1P1G1L1P1V1C1Q1D1H1E 1560
 Db 1501 G1Y1R1V1S1GER1AS1G1M1F1D1SV1L1CE1CYD1AC1AM1E1L1P1A1T1Y1V1L1A1Y1M1T1P1G1L1P1V1C1Q1D1H1E 1560
 Qy 1561 FWE1AV1FT1GL1TH1DA1F1L1S1OT1K1G1S1E1N1F1A1V1L1T1Y1Q1A1T1V1C1A1R1A1K1P1P1S1M1D1V1M1K1C1L1R1K1P 1620
 Db 1561 FWE1AV1FT1GL1TH1DA1F1L1S1OT1K1G1S1E1N1F1A1V1L1T1Y1Q1A1T1V1C1A1R1A1K1P1P1S1M1D1V1M1K1C1L1R1K1P 1620
 Qy 1621 TLV1GP1PL1Y1RL1G1V1T1N1E1V1L1TH1P1V1T1K1Y1I1AT1C1M1A1D1E1V1M1T1S1W1V1L1A1G1V1L1A1V1A1Y1C1L1A 1680
 Db 1621 TLV1GP1PL1Y1RL1G1V1T1N1E1V1L1TH1P1V1T1K1Y1I1AT1C1M1A1D1E1V1M1T1S1W1V1L1A1G1V1L1A1V1A1Y1C1L1A 1680
 Qy 1681 TGC1VC1I1GR1H1N1Q1A1V1A1P1D1K1E1V1Y1E1A1P1D1E1B1E1C1A1S1R1A1L1E1B1Q1R1A1E1M1L1K1S1K1I1Q1G1L 1740
 Db 1681 TGC1VC1I1GR1H1N1Q1A1V1A1P1D1K1E1V1Y1E1A1P1D1E1B1E1C1A1S1R1A1L1E1B1Q1R1A1E1M1L1K1S1K1I1Q1G1L 1740
 Qy 1741 QO1AS1K1O1AD1I1O1PT1Y1O1AS1P1K1V1E1Q1F1A1K1M1M1N1F1S1G1Q1Y1L1A1G1S1T1L1P1C1N1P1A1V1A1S1M1A1F1S1N1A 1800
 Db 1741 QO1AS1K1O1AD1I1O1PT1Y1O1AS1P1K1V1E1Q1F1A1K1M1M1N1F1S1G1Q1Y1L1A1G1S1T1L1P1C1N1P1A1V1A1S1M1A1F1S1N1A 1800
 Qy 1801 L1TS1PL1ST1T1L1N1L1G1M1L1S1Q1L1AP1P1A1G1T1G1V1S1G1V1A1V1S1G1L1G1K1V1L1V1D1L1A1G1Y1G 1860
 Db 1801 L1TS1PL1ST1T1L1N1L1G1M1L1S1Q1L1AP1P1A1G1T1G1V1S1G1V1A1V1S1G1L1G1K1V1L1V1D1L1A1G1Y1G 1860
 Qy 1861 AG1S1G1A1V1A1FK1M1S1E1K1S1M1E1V1N1L1L1G1L1S1P1G1L1V1G1V1C1A1L1L1R1H1V1G1B1E1A1V1Q1M1 1920
 Db 1861 AG1S1G1A1V1A1FK1M1S1E1K1S1M1E1V1N1L1L1G1L1S1P1G1L1V1G1V1C1A1L1L1R1H1V1G1B1E1A1V1Q1M1 1920
 Qy 1921 N1R1L1F1A1S1R1G1N1V1A1P1TH1Y1T1E1S1D1A1S1Q1R1Y1T1Q1L1G1S1L1T1S1L1R1L1H1M1T1D1C1P1C1G1S1M 1980
 Db 1921 N1R1L1F1A1S1R1G1N1V1A1P1TH1Y1T1E1S1D1A1S1Q1R1Y1T1Q1L1G1S1L1T1S1L1R1L1H1M1T1D1C1P1C1G1S1M 1980
 Qy 1981 L1R1D1V1M1V1C1T1L1D1E1K1M1L1S1K1L1F1K1M1P1G1L1P1V1S1C1Q1K1Y1K1G1V1A1G1T1G1M1T1R1C1P1C1G1A1N1S 2040
 Db 1981 L1R1D1V1M1V1C1T1L1D1E1K1M1L1S1K1L1F1K1M1P1G1L1P1V1S1C1Q1K1Y1K1G1V1A1G1T1G1M1T1R1C1P1C1G1A1N1S 2040
 Qy 2041 G1V1K1G1M1K1R1I1V1P1R1C1R1M1M1S1G1T1F1I1N1A1Y1T1G1P1C1L1P1A1P1N1K1F1L1M1R1S1A1E1V1E1I1R1V 2100
 Db 2041 G1V1K1G1M1K1R1I1V1P1R1C1R1M1M1S1G1T1F1I1N1A1Y1T1G1P1C1L1P1A1P1N1K1F1L1M1R1S1A1E1V1E1I1R1V 2100
 Qy 2101 G1S1Y1H1T1G1L1T1D1N1L1V1P1C1Q1L1P1S1P1E1F1S1W1D1G1Q1H1R1P1A1P1T1P1R1F1D1E1V1S1F1C1V1G1N1S1F1V1 2160
 Db 2101 G1S1Y1H1T1G1L1T1D1N1L1V1P1C1Q1L1P1S1P1E1F1S1W1D1G1Q1H1R1P1A1P1T1P1R1F1D1E1V1S1F1C1V1G1N1S1F1V1 2160
 Qy 2161 G1S1Q1L1C1D1E1P1D1V1L1M1L1D1P1S1H1T1A1R1L1A1R1G1S1P1S1E1A1S1A1S1Q1L1A1P1S1L1A1T1C1T 2220
 Db 2161 G1S1Q1L1C1D1E1P1D1V1L1M1L1D1P1S1H1T1A1R1L1A1R1G1S1P1S1E1A1S1A1S1Q1L1A1P1S1L1A1T1C1T 2220
 Qy 2221 T1G1K1A1Y1D1V1D1M1D1A1N1F1---N1G1D1V1T1R1E1S1G1K1V1V1D1S1L1D1M1V1E1R1S1L1E1S1I1S1E1Y1L 2276
 Db 2221 T1G1K1A1Y1D1V1D1M1D1A1N1F1---N1G1D1V1T1R1E1S1G1K1V1V1D1S1L1D1M1V1E1R1S1L1E1S1I1S1E1Y1L 2276
 Qy 2280 A1N1H1D1P1D1A1E1L1E1A1N1L1M1R1O1E1M1G1N1T1R1E1S1E1N1K1V1I1D1S1F1D1V1A1B1E1D1E1R1E1V1S1P1A1E1L1R 2280
 Db 2280 A1N1H1D1P1D1A1E1L1E1A1N1L1M1R1O1E1M1G1N1T1R1E1S1E1N1K1V1I1D1S1F1D1V1A1B1E1D1E1R1E1V1S1P1A1E1L1R 2280

QY	2277	PKKPEPPRLPAMAPRDYVNPVLVESKRPDYOQATYAGALAPPRKPTIPPPRRRRRTYGLS	2336
Db	2281	KSRRRAARLPWAPARDYNPLVETWKKPDYEPVNHGCPLEPPRSPVPVPPRRKRRTYVL	2340
QY	2337	EDSICDALAQQAIAISFGQPPSGDSGLSTGGAADSGSQTPPDELALSETSISSMPLE	2396
Db	2341	ESTLSTALAEALATYSFGSSSTSGITGDNTTSSSEPAEGCCPPD-----SDVSYSMPLE	2396
QY	2397	GBLGDPLEPPEQVEPQPPPGGVAAPGSDSGSWTCSF - EDSVYCCSMSSYMTGALIT	2454
Db	2397	GGPPGPDLP-----SDGSMTWSVSGADLTEDVVCSSMSYMTGALV	2436
QY	2455	PCSPEEELPIPNLSNLSLRHNKRYCTTTTSGASLRACKTFFDMQVULDSYDVLADIK	2514
Db	2437	PCAAEEQKLPINALSNSLLRRHNLYSTTSRSACQROKVTFFDRLOVYDHSYQVULKEVK	2496
QY	2515	LAASVYTAARLITMEBACOLTTPPHSARSXYGAGKEVRSLSGRAVNHISKSVKDLDEBET	2574
Db	2497	AAASRYKANLISVEBACSLTTPPHSAKSKFEGAGADVRCHAKAYAHINSVKKDLLEDSVT	2556
QY	2575	PIPTITMAKEVFCVDPYKGGKKAARLIVPDLGVRVCEKNAUYDTQKLPQAVMGASYG	2634
Db	2557	PIDITIMAKNEVFCVQPEKGGARKPARLIVPDLGVRVCEKNAUYDVVSKPLAVMGSSYG	2616
QY	2635	FOYSGAQCVEFTLLAKMAKDPKMGFSYTRCFDSTYTERDITRTEESIYRACSLPEBAHTA	2694
Db	2617	FOYSPGQVEVELVQAWMSKKTIPMGFSYDTRCFDSTYTESDITRTEBAIYQCCDLDPQARVA	2676
QY	2695	IHSILTERLYVCGPMENSKGOTCGYRCRAGSVLITTSMGNTTTCYVKKLAAACKAGIIAPT	2754
Db	2677	IKSLITERLYVGGPLTNSRGENCGRCRASGVLTTCGNTLTTCYKAPACRAAGLDDCT	2736
QY	2755	MLVCGDDLVIISGQTEDEDERNLRAFTAMTRYSAPEGDPREBYDLELITSCSSNVSV	2814
Db	2737	MLVCGDDLVIICESAGVEDAASLRAPTEAMTRYSAPEGDPREBYDLELITSCSSNVSV	2796
QY	2815	ALGQGGRRRYVLTDPDTPPIARAAMEYTRHSPVNSWLGNIQIYAPTTIMARVMTLTHFPST	2874
Db	2797	AHDGAKRKYVLTDPDTPPIARAAMEYTAARHPVNSWLGNIIMFAPTTIMARMIMLTHFPST	2856
QY	2875	LMADPTLDQNIINFEMYGAVSVSPDLDPALITERLHGLDAFSLHTYTPHELTRVSAALRKL	2934
Db	2857	LIADQLEQALNCELYGACYSIEPLDLPITQLRHGLSAFSLHSYSGEIRKRVAACTRLK	2916
QY	2935	GAPPLRAWKSFARAVALISRGGRAAVCGRYLFENMAVTKLKLTPLPEARLIDLSMFT	2994
Db	2917	GVPLPRAWHRBARASVBARLISRGGRAAICGRYLFENMAVTRTKLTPLPAAGRIDLDSMFT	2976
QY	2995	VGAGGGDIYHSVSPARPLLFGLLLFVGYGLTLLPAR 3033	
Db	2977	AGYSGGDIYHSVSHARPWFMCLLLLAAGVGYLLPWR 3015	

RESULT 11			
AAB30733 standard; protein; 3015 AA.			

AAB30733;			

02-APR-2001 (first entry)			

Amino acid sequence of chimeric Hepatitis C virus clone pH77 (p7)-J6S.			

HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine.			

Synthetic.			

OS Hepatitis C virus.			

W0200075338-A2.			

14-DEC-2000.			

02-JUN-2000; 2000MO-US015446.
04-JUN-1999; 99US-0137693P.
(USSH) US DEPT HEALTH & HUMAN SERVICES.
Yanagi M, Bukh J, Emerson SU, Purcell RH;
WPI; 2001-061728/07.
N-PSDB; AAC86648.
Nucleic acid molecule encoding human hepatitis C virus of genotype 2a for
developing vaccines, for diagnosis of hepatitis C virus and in screening
assays for identification of antiviral agents.
Disclosure; Page 150-161; 167pp; English.
AA830730-33 are encoded by chimeric cDNA clones of infectious Hepatitis C
virus (HCV). In each clone the C, E1 and E2 genes are derived from HCV
strain HC-J6ch genotype 2a. The p7 protein was derived from HCV strain HC
-J6ch or HCV strain PCV-H77C, and the NS genes were derived from HCV
strain PCV-H77C. Such HCV sequences are capable of expressing the virus
when transfected into cells. The HCV protein is useful for assaying
candidate antiviral agents for activity against HCV. Antibodies specific
for HCV polypeptide are useful in prevention and treatment of diseases
caused by HCV in animals, in particular humans. The HCV polypeptides
serve as immunogens in the development of vaccines for preventing HCV in
mammals or as antigens in diagnostic assays for detecting the presence of
HCV in biological samples. The HCV polynucleotide is also useful for
identifying cell lines capable of supporting the replication of HCV in
vivo and to produce attenuated viral strains via passage in vitro or in
vivo

Query Match	80.2%	Score 12925	DB 4	Length 3015
Best Local Similarity	78.0%	Fred. No. 0		
Matches 2371	Conservative 270	Mismatches 368	Indels 30	Gaps 4
Qy	1	MSTNPKPDRKTKRNTNRARRPODVKPFPGGQIQVGVYLLPRRGRRLGVRAIRKTSERSQPRG	60	
Dp	1	MSTNPKPDRKTKRNTNRARRPODVKPFPGGQIQVGVYLLPRRGRRLGVRAIRKTSERSQPRG	60	
Qy	61	RRQRI PKDRRSTGSGWKGPYIPWLYGNEGIGMACWLLSPGSRPSWGPNDRHRSRRNG	120	
Dp	61	RRQRI PKDRRSTGSGWKGPYIPWLYGNEGIGMACWLLSPGSRPSWGPNDRHRSRRNG	120	
Qy	121	KVIDTLTCGFADLKGVIYVPGAPLGGVBARLAHGRVLEBDGNFPTGNLPGCSFSIFLLA	180	
Dp	121	KVIDTLTCGFADLKGVIYVPGAPLGGVBARLAHGRVLEBDGNFPTGNLPGCSFSIFLLA	180	
Qy	181	LASCITTPVSAAEVKNISTGYMVTNDCTNDSITWQLOAAVLHPGCVPECEKVGNSQCIW	240	
Dp	181	LASCITTPVSAAEVKNISTGYMVTNDCTNDSITWQLOAAVLHPGCVPECEKVGNSQCIW	240	
Qy	241	PVSNVNAVQRRGALTQGLRTHIIDWVMSATLCSALYVQDLCGVNLLAOMFLVSPQHMF	300	
Dp	241	PVSNVNAVQRRGALTQGLRTHIIDWVMSATLCSALYVQDLCGVNLLAOMFLVSPQHMF	300	
Qy	301	VQDNCISIPETITGHRAMADMMNMWSPATMIILAYARVBEVIDIISGAMGVFGLA	360	
Dp	301	VQDNCISIPETITGHRAMADMMNMWSPATMIILAYARVBEVIDIISGAMGVFGLA	360	
Qy	361	YFSNQGAMAKVYVILLLLAAGVDARTHTYGGSAAOQTGALTSIFDNGPROKIQLVNTNSM	420	
Dp	361	YFSNQGAMAKVYVILLLLAAGVDARTHTYGGSAAOQTGALTSIFDNGPROKIQLVNTNSM	420	
Qy	421	HINTALTNCNDSLHTGFASLFYTHSPVSSCPPEMSACRSIEAFRYGALQYEDNVTN	480	
Dp	421	HINTALTNCNDSLHTGFASLFYTHSPVSSCPPEMSACRSIEAFRYGALQYEDNVTN	480	
Qy	481	PEDMRPYCMHYPRQGVVSAKTVCGPVYCTPSPHVVGTTDLGAPTYTMGENETDVL	540	

Db 481 PEDMBRPMHYPRROGCVVSAKTVCGPVYCFTPSPVVVGTGTRDLGAPTYTGMENETDVEI 540
QY 541 INSTRPPLGSMFGCTMWNSSGYTKTGAPPCRTBADFNASTDLCTDCEFRKHPDTTYK 600
Db 541 INSTRPPLGSMFGCTMWNSSGYTKTGAPPCRTBADFNASTDLCTDCEFRKHPDTTYK 600
QY 601 CGSGPWLPRCLIDYPRYLMHYPCITVNTYTFKIRMYGVGEHRLTAACNPTRGDRCNLE 660
Db 601 CGSGPWLPRCLIDYPRYLMHYPCITVNTYTFKIRMYGVGEHRLTAACNPTRGDRCNLE 660
QY 661 RRSQSLPLSHSTTEMAIIPCSYSDI.PALSTGLHMHONIIVQYMGYSPLATKYIYW 720
Db 661 RRSQSLPLSHSTTEMAIIPCSYSDI.PALSTGLHMHONIIVQYMGYSPLATKYIYW 720
QY 721 EWWILLFLILADARVACIMLILLOAEBALEKVIILHAASAASGNGFLYFVIFVAA 780
Db 721 EWWILLFLILADARVACIMLILLOAEBALEKVIILHAASAASGNGFLYFVIFVAA 780
QY 781 YIKGRVPLATYSLTGLMSFSLILALPOOAYAYDASVHGOIGALLVMTLEFVLTPOYK 840
Db 781 YIKGRVPLATYSLTGLMSFSLILALPOOAYAYDASVHGOIGALLVMTLEFVLTPOYK 840
QY 841 TILSRFLMCLYLLTIGEMANOEMAPPMOVRGROGIIWAVALFYRGVVPDITKMLATL 900
Db 841 TILSRFLMCLYLLTIGEMANOEMAPPMOVRGROGIIWAVALFYRGVVPDITKMLATL 900
QY 901 GPAYLLKGLATRVPEVRAHALRLMCTMARHLAGRYVQMALALGRMTGYIYDHLTPM 960
Db 901 GPAYLLKGLATRVPEVRAHALRLMCTMARHLAGRYVQMALALGRMTGYIYDHLTPM 960
QY 961 SUMAASGLDLAVANEPITFSPMEKKVYWGABETAAAGDILHGLPVSAFLGREVILGPAD 1020
Db 961 SUMAASGLDLAVANEPITFSPMEKKVYWGABETAAAGDILHGLPVSAFLGREVILGPAD 1020
QY 1021 GYTSKMSLLAPTAACOTRGLIGTIVSMTRGDRDTBOAGRIQVLTSTYOSFLGHSISG 1080
Db 1021 GYTSKMSLLAPTAACOTRGLIGTIVSMTRGDRDTBOAGRIQVLTSTYOSFLGHSISG 1080
QY 1081 VMTVYHGAAGNTLKGSRGPVTOMYSASAGDLVGMPSPGTKSLBECTCGAIDLVLVTN 1140
Db 1081 VMTVYHGAAGNTLKGSRGPVTOMYSASAGDLVGMPSPGTKSLBECTCGAIDLVLVTN 1140
QY 1141 ADVIPARRRGDXGALLSPRLSTLKSGSGGCVLCRGAHGVFRAAVCSRGVANSIDFI 1200
Db 1141 ADVIPARRRGDXGALLSPRLSTLKSGSGGCVLCRGAHGVFRAAVCSRGVANSIDFI 1200
QY 1201 PVEITDIYTRSPTEFSDNSTPRAVPOYVGYLHAPFGSGSKTKVPAYAAQGYKVLVNP 1260
Db 1201 PVEITDIYTRSPTEFSDNSTPRAVPOYVGYLHAPFGSGSKTKVPAYAAQGYKVLVNP 1260
QY 1261 SVAAATLGFAGVLSKAHGINPRTGRTVTTGAPITYSTYGFADGGCAGAYDIIICD 1320
Db 1261 SVAAATLGFAGVLSKAHGINPRTGRTVTTGAPITYSTYGFADGGCAGAYDIIICD 1320
QY 1321 ECHAVDSTIILGIGTVLDOAETAGVRLTVLATPFGSVTTPHNIEEVALQOGEIPIY 1380
Db 1321 ECHAVDSTIILGIGTVLDOAETAGVRLTVLATPFGSVTTPHNIEEVALQOGEIPIY 1380
QY 1381 GRAIPLSYIKGGRHLIFCHSKKKCDLAAALRGMGNSVAYYRGDLVSIPIYQGDVVVYA 1440
Db 1381 GRAIPLSYIKGGRHLIFCHSKKKCDLAAALRGMGNSVAYYRGDLVSIPIYQGDVVVYA 1440
QY 1441 TDALMTGYTDPDSVIDCNVAATVQVDFSLDPTFTTTOIVPODASVRSQRRGTRGL 1500
Db 1441 TDALMTGYTDPDSVIDCNVAATVQVDFSLDPTFTTTOIVPODASVRSQRRGTRGL 1500
QY 1501 GIYRYVSTGERASGMFDSVVLCECYDAGAAETLTPSETTVRLRAYENTPGLPVCODHLE 1560
Db 1501 GIYRYVSTGERASGMFDSVVLCECYDAGAAETLTPSETTVRLRAYENTPGLPVCODHLE 1560
QY 1561 FMEAVPTGLTHIDAHFLSOTKSGENFATLTAQAVVCARAKAPPSMVMKCLTRLKP 1620
Db 1561 FMEAVPTGLTHIDAHFLSOTKSGENFATLTAQAVVCARAKAPPSMVMKCLTRLKP 1620

QY 1621 TLVGPPLLYRSGVTNEVTLTHPVTXYIATCMQADLEVMSTWVLAGVLAAVAACIA 1680
Db 1621 TLVGPPLLYRSGVTNEVTLTHPVTXYIATCMQADLEVMSTWVLAGVLAAVAACIA 1680
QY 1681 TGCVCITIGRLHINQRAVAPDKVLYEAPDEMEECASRAALIEBQRIAEMLKSKIQGL 1740
Db 1681 TGCVCITIGRLHINQRAVAPDKVLYEAPDEMEECASRAALIEBQRIAEMLKSKIQGL 1740
QY 1741 OASAKOADIPTVOASMPKVOFAKXMMNFIISGIVLAGISTLPGNPAVSMMAFSA 1800
Db 1741 OASAKOADIPTVOASMPKVOFAKXMMNFIISGIVLAGISTLPGNPAVSMMAFSA 1800
QY 1801 LTPSLSTSTIILNIGMLASQIAPPAGATGVVSGVGAAGSIGHKVLVDILAGYG 1860
Db 1801 LTPSLSTSTIILNIGMLASQIAPPAGATGVVSGVGAAGSIGHKVLVDILAGYG 1860
QY 1861 AGISGALVAFKIMSGEKSPMEDVUNLPGIISPGALVVGVCATLILRRHVGPBGAVQM 1920
Db 1861 AGISGALVAFKIMSGEKSPMEDVUNLPGIISPGALVVGVCATLILRRHVGPBGAVQM 1920
QY 1921 NRIIPASRGNHVAPTHVVTESDASQRYTOLIGSLITSLRLHNMITTEDCPICGSGW 1980
Db 1921 NRIIPASRGNHVAPTHVVTESDASQRYTOLIGSLITSLRLHNMITTEDCPICGSGW 1980
QY 1981 LRDVMDVYCTIILDFKNNMLTSKLPFKMGLPEVSCOKGYKGVMACTGIMTTRCPGAMIS 2040
Db 1981 LRDVMDVYCTIILDFKNNMLTSKLPFKMGLPEVSCOKGYKGVMACTGIMTTRCPGAMIS 2040
QY 2041 GNVRLGSRKRTGPKTCNNIWOCTFPINCYTBQCYCPKAPNPKVALIWAASEVAFVQH 2100
Db 2041 GNVRLGSRKRTGPKTCNNIWOCTFPINCYTBQCYCPKAPNPKVALIWAASEVAFVQH 2100
QY 2101 GSXYITGTLTDNLKVPQOLSPSEFSPWVDGQIHRPAPTPPFPFDEVSFCVGN8FVY 2160
Db 2101 GSXYITGTLTDNLKVPQOLSPSEFSPWVDGQIHRPAPTPPFPFDEVSFCVGN8FVY 2160
QY 2161 GSQULPCDEPDPTDVLMSMLTDPSSHITAETAARLARGSPSEASSASQIAPSLRATCT 2220
Db 2161 GSQULPCDEPDPTDVLMSMLTDPSSHITAETAARLARGSPSEASSASQIAPSLRATCT 2220
QY 2221 THGKAYDVMDANLF-----MGDVTRESGKVVVLDLSLDMVEBRSDLESTSEYVL 2276
Db 2221 THGKAYDVMDANLF-----MGDVTRESGKVVVLDLSLDMVEBRSDLESTSEYVL 2276
QY 2277 PKKRPPLAPMARPDYNPPLVESMKRPDYOPATYAGCALPPRTPTPPRRRTVGIS 2336
Db 2277 PKKRPPLAPMARPDYNPPLVESMKRPDYOPATYAGCALPPRTPTPPRRRTVGIS 2336
QY 2337 EDSIGDALQOLAIKSFQGPSPSGDGLSTGAGAADSGSQTPPEDLALSETGISISMPLE 2396
Db 2337 EDSIGDALQOLAIKSFQGPSPSGDGLSTGAGAADSGSQTPPEDLALSETGISISMPLE 2396
QY 2397 GELGPDLEPPEOVEQPPPOGGVAAAPGSDSGSWSCSE--EDDSVYCCSMYSMTGALIT 2454
Db 2397 GELGPDLEPPEOVEQPPPOGGVAAAPGSDSGSWSCSE--EDDSVYCCSMYSMTGALIT 2454
QY 2455 PCSPEEKLPIPLNSLSLRHYNKYCTTTSASLRKAKVFPDRMQVDSYVDSVLKDIK 2514
Db 2455 PCSPEEKLPIPLNSLSLRHYNKYCTTTSASLRKAKVFPDRMQVDSYVDSVLKDIK 2514
QY 2515 LAASKYTABRLTMEBAQLTTPPHSARSKYGFAGKEVRSLSGRAVNHKISWMDLLEDET 2574
Db 2515 LAASKYTABRLTMEBAQLTTPPHSARSKYGFAGKEVRSLSGRAVNHKISWMDLLEDET 2574
QY 2575 PIPTTMAKNVFCVDPYTKGKKAARLIVYDPLGVRVBEKALYITOKLPAVWGAAGS 2634
Db 2575 PIPTTMAKNVFCVDPYTKGKKAARLIVYDPLGVRVBEKALYITOKLPAVWGAAGS 2634
QY 2635 FOYSPAQVBEFLKMAEKDPMGFSYDTRCDSVTVEBRDI RTEESIYRAGSLPEBAHTA 2694
Db 2635 FOYSPAQVBEFLKMAEKDPMGFSYDTRCDSVTVEBRDI RTEESIYRAGSLPEBAHTA 2694
QY 2694 FOYSPAQVBEFLKMAEKDPMGFSYDTRCDSVTVEBRDI RTEESIYRAGSLPEBAHTA 2694
Db 2694 FOYSPAQVBEFLKMAEKDPMGFSYDTRCDSVTVEBRDI RTEESIYRAGSLPEBAHTA 2694

QY	2695	HSLSLTERLYVGGPMFNKSGQFCGRRCRQASVLTTSMGNTITTCVKMLAAKAGIAPT	2753
Db	2677	IKSLTERLYVGGPILNLSRNGENCGYRCRQASVLTTSCNTILTCYIKAAACRAAGLDCT	2736
QY	2755	MLVCGDDLWVISESQTEEDERNIRAFTEAATRYASAPGDDPREYDELITSSCSNVSV	2814
Db	2737	MLVCGDDLWVISESAGVOEDPAASIRAFTEAATRYASAPGDDPREYDELITSSCSNVSV	2796
QY	2815	ALGPOGRRRYYLTEDPTTPIARAAMEYTRVSPVSWLGNIIQVAPTIWAWVLMTTFPSI	2874
Db	2797	AHDAGKRVYYLTREDPTTPIARAAMEYTRHVPVSWLGNIIIMPFTLMARNIIMTTFPSV	2856
QY	2875	LMADDTLDQNLNFMENYGVAVSVSPLDLPATIERLHGLDQFSGLTYTTPHELTRVASALRKL	2934
Db	2857	LIARDQLEQALNCEIYGACYSIEPLDLPPIQLRHGSAFSLHSYSPGEINRVAACLRL	2916
QY	2935	GAPPLRWKSKAARAVRASLISRGGRAVCGYYLFWMAVTKLTKTPLPEARLLDSSNFT	2994
Db	2917	GVAPPLRWKMRHARAVRARLLSRGERAAICGYLFLPMAWARTKLKTLPIAAGRLLDSSWFT	2976
QY	2995	VGAGGGDIYHSVSRARPLLFGLLLLFVYGFLPLPAR	3033
Db	2977	AGVSGGDIYHSVSHARPRWFCLLLLAAGVGIYLLPAR	3015

RESULT 12
 AAB30731
 ID AAB30731 standard; protein; 3015 AA.
 XX
 AC AAB30731;
 XX
 DT 02-APR-2001 (first entry)
 XX
 DE Amino acid sequence of chimeric Hepatitis C virus clone pH77(p7)CV-J6S.
 XX
 KM HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine.
 XX
 OS Synthetic.
 OS Hepatitis C virus.
 XX
 MN MO200075338-A2.
 XX
 PD 14-DEC-2000.
 XX
 PF 02-JUN-2000; 2000MO-US015446.
 XX
 PR 04-JUN-1999; 99US-0137693P.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Yanagi M, Bukh J, Emerson SU, Purcell RH;
 DR WPI; 2001-061728/07.
 DR N-PSDB; AAC86646.
 XX
 PT Nucleic acid molecule encoding human hepatitis C virus of genotype 2a for
 PT developing vaccines; for diagnosis of hepatitis C virus and in screening
 PT assays for identification of antiviral agents.
 XX
 PS Disclosure; Page 119-130; 167pp; English.
 XX
 AAB30730-33 are encoded by chimeric cDNA clones of infectious Hepatitis C
 CC virus (HCV). In each clone the C, E1 and E2 genes are derived from HCV
 CC strain HC-J6CH genotype 2a. The p7 protein was derived from HCV strain HC-
 CC J6CH or HCV strain PCV-H77C, and the NS genes were derived from HCV
 CC strain PCV-H77C. Such HCV sequences are capable of expressing the virus
 CC when transfected into cells. The HCV protein is useful for assaying
 CC candidate antiviral agents for activity against HCV. Antibodies specific
 CC for HCV polypeptide are useful in prevention and treatment of diseases
 CC caused by HCV in animals, in particular humans. The HCV polypeptides
 CC serve as immunogens in the development of vaccines for preventing HCV in
 CC mammals or as antigens in diagnostic assays for detecting the presence of
 CC HCV in biological samples. The HCV polynucleotide is also useful for

CC	Identifying cell lines capable of supporting the replication of HCV in
CC	vitro and to produce attenuated viral strains via passage in vitro or in
CC	vivo
CC	XX
CC	Seq
	Sequence 3015 AA;
	Query Match 80.2%; Score 12925; DB 4; Length 3015;
	Best Local Similarity 78.0%; Pred. No. 0;
	Matches 2371; Conservative 270; Mismatches 368; Indels 30; Gaps 4
Qy	1 MSTNPKPQKRTKRNTRRRPODVKEFGGQIVGGVYLIPRRGPRLGATRAKTKTSERSQPRG 60
Db	1 MSTNPKPQKRTKRNTRRRPODVKEFGGQIVGGVYLIPRRGPRLGATRAKTKTSERSQPRG 60
Qy	61 RRQPIPKDRRSTGSMGKRGYPWPLYNGEGLCAACMLSPRSRSPSGMNDPRHRNRNG 120
Db	61 RRQPIPKDRRSTGSMGKRGYPWPLYNGEGLCAACMLSPRSRSPSGMNDPRHRNRNG 120
Qy	121 KVITDTLTCGPFADLMGVIPIVGAAPLGGVBARALAHGRVLEEDGNFPAATGNLPGCSFSIFLLA 180
Db	121 KVITDTLTCGPFADLMGVIPIVGAAPLGGVBARALAHGRVLEEDGNFPAATGNLPGCSFSIFLLA 180
Qy	181 LLSCITTPVSAEAVKNISTGYWNTDCTNDSITWOLQAAVILHPCVPCERKYNASQCI 240
Db	181 LLSCITTPVSAEAVKNISTGYWNTDCTNDSITWOLQAAVILHPCVPCERKYNASQCI 240
Qy	241 PVSPPVNAVQRGALITQGLRTHIDMWYMSATLCSALVYDILCGGVMLAAOMFVSPQHMF 300
Db	241 PVSPPVNAVQRGALITQGLRTHIDMWYMSATLCSALVYDILCGGVMLAAOMFVSPQHMF 300
Qy	301 VQDCNCSITPGTITGHRWAMDMMNMSPTATMILAYANRVEVILIISGAHGVNFGLA 360
Db	301 VQDCNCSITPGTITGHRWAMDMMNMSPTATMILAYANRVEVILIISGAHGVNFGLA 360
Qy	361 YFSNQGAMAKVYVILLILAAGVDARHTYTGSSAAQTGRLTSLFDMGPRROKIQLVNTGSM 420
Db	361 YFSNQGAMAKVYVILLILAAGVDARHTYTGSSAAQTGRLTSLFDMGPRROKIQLVNTGSM 420
Qy	421 HINRTALNCNDSLTGTFTASLFYTHSPFSSGCPSEMSACRSITAEFRVGVGALQYEDNVTN 480
Db	421 HINRTALNCNDSLTGTFTASLFYTHSPFSSGCPSEMSACRSITAEFRVGVGALQYEDNVTN 480
Qy	481 PEDMRPYCMWIPRQCGVSAKTVCGPYCTPSPVVVGTDRDLGAPTYTWGENETDVL 540
Db	481 PEDMRPYCMWIPRQCGVSAKTVCGPYCTPSPVVVGTDRDLGAPTYTWGENETDVL 540
Qy	541 LNSRPRPLGSPFGCTMMNMSGYTTKCGAPPCRTTRADFNASTDLCLPTDPCRHHPTTYLK 600
Db	541 LNSRPRPLGSPFGCTMMNMSGYTTKCGAPPCRTTRADFNASTDLCLPTDPCRHHPTTYLK 600
Qy	601 CGSGPMLIPRCLIDYPRYLMMYPCTVANTIKIRIMYVGVGEHRLTAACNFTRGDRCNLED 660
Db	601 CGSGPMLIPRCLIDYPRYLMMYPCTVANTIKIRIMYVGVGEHRLTAACNFTRGDRCNLED 660
Qy	661 RDRQOLSPILHSTTEMALIPCSYSDLPALSTGLLHMHONIVDVQMYGSPALTKTIVRM 720
Db	661 RDRQOLSPILHSTTEMALIPCSYSDLPALSTGLLHMHONIVDVQMYGSPALTKTIVRM 720
Qy	721 EMVILLFLLADAVACACMLMILLGQAEALIEKLVIILHAASAAACNGFLYVIFPVAAM 780
Db	721 EMVILLFLLADAVACACMLMILLGQAEALIEKLVIILHAASAAACNGFLYVIFPVAAM 780
Qy	781 YIKRRVVPPLATYSLTGLMSPSULLIALPQAAVADASYHGOIGALLVMITLFTLTPEYK 840
Db	781 YIKRRVVPPLATYSLTGLMSPSULLIALPQAAVADASYHGOIGALLVMITLFTLTPEYK 840
Qy	841 TLISRFLLMILCYLLTLTGAAVQEAAPMQVQVGRGDIIMAVAIFFPGVAFDITKMLLAUL 900
Db	841 RYISWCMWMLQYPLTRVEDAQHLWVPRPLANVGSDAVAILMNCVHPFLVFDITKLLIAIF 900
Qy	901 GPAYLLKGLALTRVYFVRAHALLMCTMARHLAGRVYOMALLALGRMTGYIYDHLTPM 960
Db	901 GPAYLLKGLALTRVYFVRAHALLMCTMARHLAGRVYOMALLALGRMTGYIYDHLTPM 960

QY 961 SPWAAGLRLDAVAVEPIIFSPMEKKYIVWGAETAAAGDILHGLPYSAARLGREVLIGPAD 1020
 Db 961 RMAHNGLDLDAVAVEPVFSPMETKILTMGADTAAAGDILINLPISARSGEILLGPAD 1020
 QY 1021 GYTSKMSLIAPITAYAOOTRGLGTTIVSMGRDXTBOAGETQVLSVTVTQSFLGTSISG 1080
 Db 1021 GWSKSMRLIAPITAYAOOTRGLGTTISLGRDKNOVEGEVOIYSTATQTFALACIN 1080
 QY 1081 VMTVYHAGANKTLASRGVPTOMYSSABGDLVGMPSPGTKSLBECTGADVLYVTNN 1140
 Db 1081 VCTVYHAGACTRTIASPKGFVIMYTNVDDLVGMFAPQGSRLPCTGSSDLVLYVTH 1140
 QY 1141 ADVIPARRRGDKGALLSPRLSTLKGSSGPLYCRGNAVGFRAAVCSRGVAKSIDPI 1200
 Db 1141 ADVIPARRRGDSKSLSPRLSITLKGSSGPLYCPAGNAVGLFRAAVCTCRGAKXVDPI 1200
 QY 1201 PVETLDIVTRSPFFSDNSTPPAVPQTYQYGLHAPPGSGKSTVPVAYAAQYKVLVNP 1260
 Db 1201 PVENLGTTRKSPVFTDNSSPPAVPQSFQVAMHLAPGSGKSTKVPAAVYAAQYKVLVNP 1260
 QY 1261 SVAAATGFGAYLSKAHGINPIRTGYRTVTGAPITYSTYKFLADGGAGAYDIIICD 1320
 Db 1261 SVAAATGFGAYMSKAGVDPNIRTKYRTITTSPIITSTYKFLADGGAGAYDIIICD 1320
 QY 1321 ECHAVDSTTILIGIVLDQAEAGVRLTVLATATPPGSVTPPHNIEEVALGQEBRIPFY 1380
 Db 1321 ECHSTDATSILIGIVLDQAEAGVRLTVLATATPPGSVTVSHPNIEEVALSTGSIPIFY 1380
 QY 1381 GRAIPLSYKGGHILFCHSKKKCDLAAALRMGINSVAYYRGDLVSYITPQGDVVVVA 1440
 Db 1381 GKAPILEVIKGGHILFCHSKKKCDLAAALVGINAVAYYRGDLVSYITPQGDVVVVS 1440
 QY 1441 TDALMTGYGDPSPSYDCNVAVTVQVDFSLDPFTTITQIVPDADVRSORRGTRGRGL 1500
 Db 1441 TDALMTGYGDPSPSYDCNTCVTQVDFSLDPFTTITQIVPDADVRSORRGTRGRGP 1500
 QY 1501 GIYRYVSTGERASGMDSVVLCECYDAGAAWELTSEETTVRLRAYFNTPGLPVCCDHL 1560
 Db 1501 GIYRFVAPGERSPGMDSVVLCECYDAGCAWELTFAETTVRAIYMTNPGLPVCCDHL 1560
 QY 1561 FMEAVTGLTHIDAHFLSQTKSGENFAYITAYQATVCARAKAPPSMDVMKCLTRLP 1620
 Db 1561 FMEGVFTGLTHIDAHFLSQTKSGENFPYLAQATVCARAQAPPSMDVMKCLTRLP 1620
 QY 1621 TLVGPRLLYRGSVYNEVTLTHPVTKYIATCMQADLEWMTWVLAAGVLAAYCYLA 1680
 Db 1621 TLHGPTPLLYRGLAVQNEVTLTHPTKYIATCMQADLEWMTWVLAAGVLAAYCYLS 1680
 QY 1681 TGCVCITIGRLHINQRAVVAPEKEVLYEAPDEMEECASRAALIEBQRIABMLKSKIQLL 1740
 Db 1681 TGCVVIVGRIYLSGKRALIPDRREVLYQEPDEMEECQHLPYIBQMMLLBQFOKALGILL 1740
 QY 1741 QOASKAQADIQPTVOASMPKVBQFMAKMMNFISGIQYIAGLSTLPQNPVASMMAFSAA 1800
 Db 1741 QTSRRAEAVITPAVQNTWQKLEVFMAKMMNFISGIQYIAGLSTLPQNPVASTMAFSA 1800
 QY 1801 LNSPLSTSTTILNIIIGMLASQIAPRAGATGVVSGVGAAGSITGDKVLDVLIAGV 1860
 Db 1801 VTSPLTGTQTLNIIIGMVAQOLAPGAATAVGAGAAIGSVGLGVLDVLIAGVG 1860
 QY 1861 AGISGALVAFKINSGEKSPMEDVNNLFGILSPGALVVGVCIAIIRRHVGPBGAGVQNM 1920
 Db 1861 AGVAGALVAFKINSGEKSPMEDVNNLFGILSPGALVVGVCIAIIRRHVGPBGAGVQNM 1920
 QY 1921 NRIIAFASRGNAVAPTHYVTESDASQRTVQLGSLTITSILRLRLHMTIEDCPICGSGM 1980
 Db 1921 NRIIAFASRGNAVAPTHYVTESDAAARVTAIISLTVQLRLRLHOMISECTTPCGSGM 1980
 QY 1981 LRDVMTVWCTLLTDFKNWLSKLFPMKPGLPFVSCQKGYGVAGGIMTTRPCGANIS 2040
 Db 1981 LRDVMTVWCTLLTDFKNWLSKLFPMKPGLPFVSCQKGYGVAGGIMTTRCHCAEAIT 2040

QY 2041 GNVRLGSRNRTGPKTCNNIWOCTFPINCYTEGOCVPEKPAENFKVAILNREVAEVAETOH 2100
 Db 2041 GHVKNGTWRIYIGPRICRMMMSGTFPINAYTTGSPCTPLAPNPKFALMVSABEYIEIRV 2100
 QY 2101 GSYHTTGLTIDNLKVPQQLSPSPFSWVDQVQHIFRAPTKPFRFSDVSVFVGLNSFVY 2160
 Db 2101 GDFVHSGMTIDNLKPCQIIBSPFEFTLDGRLIRFAPPCPPLRREVSFFVGLHEFV 2160
 QY 2161 GSQPCDEPPTDVVMSMLTDPSSHITAEARRLRSGSPSASASQSLAPSRLACT 2220
 Db 2161 GSQPCDEPPTDVAVVLSMLTDPSSHITAEARRLRSGSPSASASQSLAPSRLACT 2220
 QY 2221 THGKAYDVDMVDANLF---WGDVTRIESGKVVVLDLDPWEERSDLEPSISEYML 2276
 Db 2221 ANHDSPEDELIEANLIMRQEMOGNITRVESEKVVILDSFPLVABEDEREVSVAEILR 2280
 QY 2277 PKKRPPPLPAMARDYNNPLVESMKRDYQATYAGCALPPPKTTPPPRRRTYGLS 2336
 Db 2281 KSRREARALPVWARDYNNPLVETWKRPDYEPVHGCPLPPRSPVPVPPPKKTVLT 2340
 QY 2337 EDSIGDALQOLAIFSGQPPSPGDSGLSTGAGAADSQTPPEDELALSETGISAMPPL 2396
 Db 2341 ESTLSTALAEALATKSGSSSTISGITGNDTTSSSEAPSGCPEP---SDVESYSSMPLE 2396
 QY 2397 GELGDPDLEPEQVEBPQPPQGVAAFGSDSGSWSTCSB--EDDSVYCCSMYSWTGALIT 2454
 Db 2397 GEPGDPDL-----SDGSWSVTSAGADTBEDVCCSMYSWTGALVY 2436
 QY 2455 PCSPEBEKLPINPLNSILRTHNKYCTTTGSASLRAKKVTFDDNQVLDSTYDSVLKXK 2514
 Db 2437 PCAABEOKLPINALSNSILRHNLYSTTSRAACQKQKVFDRQVLDVSHQVPLKVK 2496
 QY 2515 LAASRYTARLITMSEACQLTPPHSARSKYGFAPKSEVRSLSGRAVHIRKSVMKDLEDET 2574
 Db 2497 AAASKYKANLISVEBACSLTPPHSAKSKFGYAKQVRCARAKVAHNSVWKDLEDET 2556
 QY 2575 PIPRTIMAKNEVYCDPTPKGKKARLIVPDLGVYCEKMAIYDITOKLPOAVNGASYG 2634
 Db 2557 PIDTTIMAKNEVYCVQPEKGGKPARLIVPDLGVYCEKMAIYDVVSKLPLAVNGSSYG 2616
 QY 2635 FOYSPAQVVEPLTKAMAKKQPMGFSYDTRCDSYVTERDRIETESIYRAGSLPREAHTA 2694
 Db 2617 FOYSPAQVVEPLVQAMKSKTTPMGFSYDTRCDSYVTERDRIETESDIYQCCDLDOARVA 2676
 QY 2695 IHSLETERLYVGGPMNSKQOTCYGRCRASGVLITSMGNTITCYVKALAACRAAGIIAPT 2754
 Db 2677 IKSLETERLYVGGPLTNSGENCGYRCASGVLITSCGNTITCYIKAAACRAAGLOCT 2736
 QY 2735 MLVCGDDLVVISGQTEDEBERNLRAFTBAMTRYSAAPQDPPRPBYDELEITSCSSNVY 2814
 Db 2737 MLVCGDDLVVICESAGVEDAASLRAFTBAMTRYSAAPQDPPRPBYDELEITSCSSNVY 2796
 QY 2815 ALPGQRRRYITTRPPTPIARAAMEYVSHSVNWSLNIIOYATTIARWMLMTHFSI 2874
 Db 2797 AHDGAKRYVYITRDPPTPLARAAMEYVSHSVNWSLNIIMFATPLIARMLMTHFSV 2856
 QY 2875 LMAODTLOONLNFEMGYAVSVPLDPAIIEIRLGLDAFSLHTYTPHELTRFVASRL 2934
 Db 2857 LTAHQDEBALNCEIYAGACTSIEPLDPEIIOKRLGLBAFSLHSISPGSINRYAACIKRL 2916
 QY 2935 GAPPLRAMKSRAAVRASLISRGRAVCGRYLFNMAVYTKLKLTPLEBARLLDSSWFT 2994
 Db 2917 GVPPLRAWRHRRARSVRARLSRGRAALCGKLFNMAVRYTKLKLPIAAGRLDSSWFT 2976
 QY 2995 VAGGGDIYHSVSRARPLLLFGLLLVGVGLFLLLPAR 3033
 Db 2977 AGYSGDDIYHSVSHARPRMFWFCLLLLAAGVGIYLLPR 3015

RESULT 13

ABG32459

ID ABG32459 standard; protein; 3010 AA.

XX

AC	ABG32459;	
XX	15-NOV-2002 (first entry)	
XX	Hepatitis C virus Con 1 isolate polypotein mutant #8.	
DE	HCV, Con 1; adaptive mutation; liver failure; cirrhosis; mutant; muten;	
KW	hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;	
XX	internal ribosome entry site; IRES; NS5A; HCV replication; polypotein.	
OS	Hepatitis C virus.	
OS	Synthetic.	
FT	Key	Location/Qualifiers
FT	Misc-difference 2198	
XX	/note= "Wild-type Leu substituted by Ser"	
PN	MO200259321-A2.	
PD	01-AUG-2002.	
XX	16-JAN-2002; 2002WO-EP000526.	
XX	23-JAN-2001; 2001US-0263479P.	
XX	(RICE-) 1ST RICEACHE BIOL MOLECOLARE ANGELETTI.	
PA	De Francesco R, Migliaccio G, Paonessa G;	
PI	WPI; 2002-599793/64.	
DR	New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV	
PT	NS5 encoding region, or encephalomyocarditis virus (EMCV) internal	
PT	ribosome entry site (IRES) region, useful in studying HCV replication and	
PT	expression.	
XX		
PS	Claim 1; Page; 63pp; English.	
XX		
CC	The invention relates to nucleic acid molecules comprising altered HCV	
CC	NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)	
CC	internal ribosome entry site (IRES) region coding for one or more NS3,	
CC	NS5A, or EMCV IRES mutations, respectively. The location of the mutations	
CC	are detailed in the specification. Also included are (1) an expression	
CC	vector comprising a nucleotide sequence coding for the altered nucleic	
CC	acids, which is transcriptionally coupled to an exogenous promoter; (2) a	
CC	recombinant cell human hepatoma cell comprising the altered nucleic acids	
CC	; (3) a recombinant cell produced by introducing into a human hepatoma	
CC	cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)	
CC	replicon enhanced cell or which containing a functional HCV replicon; (5)	
CC	an HCV replicon enhanced cells made in the method; and (6) measuring the	
CC	ability of a compound to affect HCV activity. The HCV replicons and HCV	
CC	replicon enhanced cells are useful in studying HCV replication and	
CC	expression, and HCV and host cell interactions, producing HCV RNA and	
CC	proteins, and providing a system for measuring the ability of a compound	
CC	to modulate one or more HCV activities e.g. to discover drugs which may	
CC	treat HCV mediated diseases such as liver failure, cirrhosis and	
CC	hepatocellular carcinoma. The present sequence is the HCV replicon Con 1	
CC	polypotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A	
CC	and NS5B proteins), NS5A mutant of the invention. Note: The present	
CC	sequence is not shown in the specification but was created by the indexer	
CC	using the HCV sequence appearing as ABG32451 and the information in claim	
CC	1	
XX		
XX	Sequence 3010 AA;	
XX		
QY	Query Match 74.7%; Score 12034.5; DB 5; Length 3010;	
	Best Local Similarity 71.8%; Pred. No. 0;	
DB	Matches 2183; Conservative 352; Mismatches 469; Indels 35; Gaps 7	
	1 MSTNKPQPKTKRNTNRPPDYKFPGGGIVGVVLLPRRGLGVRAARKTSERSQPRG 60	
	1 MSTNKPQPKTKRNTNRPPDYKFPGGGIVGVVLLPRRGLGVRAARKTSERSQPRG 60	

QY	61	RROPIPKORRSTGKSGKPGYPMPLYGNEGICJMAAGMLISPRGSRPSWGPNDPRHRNRVG	120
Db	61	RROPIPKARQCEGRMAQPGYPMPLYGNEGICJMAAGMLISPRGSRPSWGPNDPRHRNRVG	120
QY	121	KVIDITTCGFADLMGXIPVGAIPGCVARALAHGRLVEDQVNFATGMLPGCSFIFILLA	180
Db	121	KVIDITTCGFADLMGXIPVGAIPGAGARALAHGRLVEDQVNFATGMLPGCSFIFILLA	180
QY	181	LLSCITTPVSAAEVANIISTGYMTNDCINDSIITVQLAVALVHPVCCEKEKGNASQCI	240
Db	181	LLSCITTPVSAAEVARNISGVYHVTINDCSNBIIVYEAADMINTPBCVCEVKNSSRCMV	240
QY	241	PVSPVNAVOPFGALTQGLRTHIDMVVNSATLCSALVYGDLCGGVMLAAQMFIVSPQHWF	300
Db	241	ALPTETLAARNASVPTTTRIRHVDLVCGAALCSAMVYVDLCSGVFLVQLFFSPRRHET	300
QY	301	VODNCSTIYPBTIIGHRAWMNNMNSPTAMILAYARVBEVITDIISGAWGMPCGA	360
Db	301	VODNCSTIYPGHVIGHRAWMNNMNSPTALVVSQILRIQAVVDVMAAGHWGLAGLA	360
QY	361	YFSMOGAAKVVVLLLLAAGVDARTHTYGGSAAQTCGTLSTLFDMPGPROKIQLVNTNGSM	420
Db	361	YYSVWGNMAKULIYMLFAGVDGCTYTYGTMAKNTLGITSLFBSGSSQKIQLVNTNGSM	420
QY	421	HINRTALNCNDSLHTGFTASLFTYHSFNSSGCCPERMSACRSIEAFRVGMAQYEDVNTN	480
Db	421	HINRTALNCNDSLNTGFLALFYVHKFNSSGCCPERMSACSPIDAFQCMGPITY--NESH	478
QY	481	PEDMRPYWNIHPPROCGVVSATYCGPYCTSPSPVNVGTTDLGAPRTYTGENTDVL	540
Db	479	SSDRPYWNIHAPPCGIVPAQAQVCGPYCTSPSVNVGTTDRGVPTYSWGENETDVL	538
QY	541	LNSTRPPLGSGFSGCMNNSSGYTTCGAPPCPRTBADAFASTDLCEPTCCFRHGPPTYLK	600
Db	539	LNSTRPPOGAMFGCTMNNSTGFTKTCGSPPCNIGIGIKNT--LTCPTDCFRHGPATYTK	596
QY	601	CGSGFWLPRCLAIYVYRLMHPCTVNTTIEKIRMYGVGEHRILTAACNFTGRBCNLED	660
Db	597	CGSGFWLPRCLVHPYRLMHPCTVNTTIEKIRMYGVGEHRILEAACNMTGRGERCNLED	656
QY	661	RDRSOLSPHLSTHEMALIPCSYSDDLPAALSTGLLHLHONIVDQPMGLSPALTYIYRW	720
Db	657	RDRSELSPHLSTHEMOVLPCSFPTLPAALSTGLLHLHONVVDVOYLIGISAVSFAIKW	716
QY	721	EMVILLFILLADAVVCACTMMLILLGQAEALBKULIHAASASCNCFVLFVFPVAM	780
Db	717	EYVILLFILLADAVCACLMMLILIAQAEALBNLVYNAASVAGAGILSLVFFCAN	776
QY	781	YIKRVPVPLATYSTGLMSFSLILLALPOQAVAYADSVHGOIGALLVMTLFTLLPGYK	840
Db	777	YIKRIVPGAAVYALYGVWPLILLILLALPRAYVANDREMAASCQGAIVPGILLILTLSPHYK	836
QY	841	TLISRPILMTLCYLLTLGEAMVOEAPPMQVRCGRDGIIMAVAIIFPGVFPDITKULAVL	900
Db	837	LFELRLIMVLOQYFTRIEAHLQVMIPLPNTVGGSDAVILLITCAIHPELIFTITKILLAIL	896
QY	901	GPAILLKGALIRVYFVYRAHALLMCTMAHILAGRVYOMALLLKGRTGYIYDHLPM	960
Db	897	GPLVNLQIGITKVYFPAHAGLIRACMLVRYVAGGHVYOMALLKUALITGYIYDHLPL	956
QY	961	SDMAASGLRDAAVVEPIIFSPMEKCVLWMAETPAACGDIHLGLPVSARLGREVLLGPAD	1020
Db	957	RDMAHAGRDLAAVVEPIVPSDMETKVI TWGADTPAACGDIHLGLPVSARGREIHLGPAD	1012
QY	1021	GYTSKGMSLAPITAYAOQTRGLLGTIVSMTGSDKTEQAGEIOVLSTVTSOFLCTSIISG	1080
Db	1017	SLGEOGRMLLAPITAYAOQTRGLLGTIITSITLGRDRNOVEGEVQVSTATOSPLATCVNG	1076
QY	1081	VLMTVYHAGAGKTLTAGSGPYTOMYSASBGLVQMPSPGKSLIEPCGCAVVDLYLYRN	1144
Db	1077	VCMTVYHAGAGSKTLTAGSGPYTQWYTNVDDVLVWQAPPGARSILPCTCGSSDLYLYTRH	1133
QY	1141	ADVLPARRRGKRGALLSPRLSTYTKSGSGGPVLCPRHAGAVFRPAAACSSGVAKSIDFI	1200

Db	1137	ADVPVRRRGSGRGLSPRPVSYLKGSGGELLCPSHAAVGIFFAAACTGTGVAKAADV	1196
QY	1201	PVETLDTITRBPPTFSNDSTPPAPVOTYOVGYLHAPTGSKSTKYVAAAGYKVLVNP	1260
Db	1197	FVESMETTRBSGVFTDNSSPAVPOTFQVHLHAPTGSKSTKYPAALAGYKVLVNP	1256
QY	1261	SVAATLFGAYLSKAGINPNIRIGVRTVTGAPITYSYTGKFLADGGCAGAYDIIICD	1320
Db	1257	SVAATLFGCAVMSKAHGIDPNIRIGVRTITGAPITYSYTGKFLADGGCAGAYDIIICD	1316
QY	1321	ECHAVDSTIIIGIGVULOAEFAGRLVLVATAPRGSVLTTPHPIEBVALGQGEIIPFY	1380
Db	1317	ECHSDSTSTIIIGIGVULOAEFAGRLVLVATAPRGSVLTTPHPIEBVALSSTGBIIPFY	1376
QY	1381	GRAIPLSYIKGGRHLIFCHSKKCCBELAAALRGMLNSVAAYRGIDVSIPTQGDVVVVA	1440
Db	1377	GKAIPFIETIKGGRHLIFCHSKKCCBELAAKLSGLGNVAAYRGIDVSIPTSGDVIYVA	1436
QY	1441	TDALMTGYTDFDSVIDCNVAVTQVDPFSLDPTFTTITQIIVPOAVSHSQRRGTRGRL	1500
Db	1437	TDALMTGFTGDFDSVIDCNVCTQVTFVDPFSLDPTFTFIETTTVPQAVSHSQRRGTRGRL	1496
QY	1501	GIRYRVSTGERASGNFDSVULCECCDAGAAYELTSPSTYRLRAYFMTBGLPVCODHLE	1560
Db	1497	GIRYRVSTGERSGNFDSVULCECCDACAWEILTSPSTYRLRAYMTBGLPVCODHLE	1556
QY	1561	FWEAVFTGLTHDAFPLSOTKOSGENFAYLTAQVATVCAKAAKAPPSPDVMKCLTRLKP	1620
Db	1557	FWESVFTGLTHDAFPLSOTKQAGNFAYLVAQVATVCAKAAKAPPSPDVMKCLTRLKP	1616
QY	1621	TLVGFPTLLYRLGSVTNEVLTLPVTKYIATCMQADLEVMTSTWVLAGVLAAYAYCLA	1680
Db	1617	TLHGFTPLLRYLGAVONEVLTTHPTTKYIMACMSADLEVMVSTWVLVGVVLAALAAVCLT	1676
QY	1681	TGCVCILIGRLHNOCAVVAPEKCVLYEAPDEMECASPAALIEBQRIAEMLKSIQGL	1740
Db	1677	TGSVAVIVARIIISGRPAIIPDEEVLYREDEMEBCASHPIYIQQMOQLAEOPKOKAIGLL	1736
QY	1741	COASKOADIOPTVOASMPKVCQPAKAMNNFISGIQVLAGLSTPGNPAVAMMAPEASA	1800
Db	1737	QTATQAEAAAPVAVESKRTLEFAFAKAMNNFISGIQVLAGLSTPGNPAVLAASMAFAS	1796
QY	1801	LTSPLSTSTIIILNIGMLASQIAPPAAGATGVVSGLVGAVSGIGLKVLYDILAGYG	1860
Db	1797	ITSPILTQHTLIFNLIGGVAAQIAPPSAASFVAGIAGAAVSGIGLKVLYDILAGYG	1856
QY	1861	AGISALVAFKIMSEKSPMEVYNLLIGLISPALVGVCAALLRHHVPGEGAVQM	1920
Db	1857	AGVAAALVAFKVMSEMPSTEDLVMLPALISPGLVGVCAALLRHHVPGEGAVQM	1916
QY	1921	NRLIAPFARGNVAPETHVYTESDAQRYTOLLGSLTITSLRRLNMTITEOCPIPCGGSS	1980
Db	1917	NRLIAPFARGNVAPETHVYTESDAARATQIILSITTLQIKLRLHQMINEOCSTPCSSSS	1976
QY	1981	LADVAVWCTIILTDEKMWLTSKLFPKMGPLPFSVCQKYGKVMAGTGMITTRPCGANIS	2040
Db	1977	LADVAVMCTIYLTDFKMWLQSKLPLRPGVPFSCQRYGKVMGDDGIMQTTCCGAQIT	2036
QY	2041	GNVRLGSKRITGPCKTQNMIMQITPINCYSBQCPKPAEPFKALIMVAASEVAEVQH	2100
Db	2037	GHVXKGSKRIVGPRKTCNSMTQITFIPINAYTGPCTBPAPNYSRLIMVAAEEVYEVFRV	2096
QY	2101	GSYHYITGTITNLIKVPCLPPEPFSFWDGQIHRFAPTPKPPFRDRESVCVGLNSPVV	2160
Db	2097	GFHYVITGMTIDNVKCPQVPAPEPFEYVDGRLHRIVAPACKPLIREVTLVAGNOVLV	2156
QY	2161	GSQLPCEDEPDTVLMSTMUTPSHTTAETAAARLARSGSPSSASSASQLSAPSLRATCT	2220
Db	2157	GSQLPCEPEPDVAVLTSMUTDPSHTTAETAKRLARSGSPSSASSASQLSAPSLRATCT	2216
QY	2221	THGKAYDVDMVDANTF---MGGDVTRIESGKVVVLDSDMWEERSDLEPSIPEEYML	2276

Accession	Protein	Gene	Species	Length (aa)	Source
2217	TRHDSPADADLTLEANLNRQEMGNIITVESSENKVIILDSFEPLQAEEDEREVSVAEILR	TRHDS	Human	2276	GenBank
2277	PKKFPFALPMAPADVNPPLVSWKSKDYQATVACALPPPKPTPTPPRRRRITGLS	PKKFP	Human	2336	GenBank
2277	RSRRFPFAMPIMAPPDVNPPLLSBWMKPDVPPVPHOCPLPPAPAKPLPPRRRRKTVLS	RSRRFP	Human	2336	GenBank
2337	EDSISGDALIQOLATISFGQPPPSG-DSGLSGTGACAADSGSQTPPELALSTGSISSNPPL	EDSISG	Human	2395	GenBank
2337	ESTVSSALAEATATPTFGSSSESSAADSGTATV----SPDPDSDDGDGSDVESYSNPPL	ESTVSS	Human	2391	GenBank
2396	EGEIGDPPLEBPEQVPEQPQCGVAALAGSDGSGWSTGSEE-DGSVCCSMSSYMTGALIT	EGEIGD	Human	2454	GenBank
2392	EGEPGDDPL-----SDGSNSTVSEESSEPVCCSMSTYTGALIT	EGEPGD	Human	2431	GenBank
2455	PCSPBEERKLPINPLNSILRYHNKYCTTTKSASLRKAKVTFDRMQVLDSEYSDVLDIRK	PCSPBE	Human	2514	GenBank
2432	PCAAETKLPINALNSGLRHHNLVYATTSASALRQKVTFDRLQVLDHRYRLAKMK	PCAAET	Human	2491	GenBank
2515	LAASKVTARLLTMEACQLTPPHSARKYGFGAKEVSLSGRAVNHKISWKDLLEBSET	LAASKV	Human	2574	GenBank
2492	AKASTVYKAKLLSVEEACQLTPPHSARSKFVGAGADVNLSSKAVNHISRWKDLLEBSET	AKASTV	Human	2551	GenBank
2575	PIPTTIAMKNNFVCDVPTKGGKKAARLLVYDGLGVRCSEKVALVDIQOKLPOAVMGASYG	PIPTTI	Human	2634	GenBank
2552	PIDTTIAMKNNFVCGQPEKGGKKAARLLVFPDGLVRCSEKVALVDVSTLPOAVMGASYG	PIDTTI	Human	2611	GenBank
2635	FQYSPAGQVEBELKAMAEKKDPMGFSYDTRCFDSVTYERDITREESLYRACSLPEAHATA	FQYSPA	Human	2694	GenBank
2612	FQYSPGQVEBELVANAKAKKCPMGFAVDTRCFDSVTYENDIRVESIYQCDDLAPEARQA	FQYSPG	Human	2671	GenBank
2695	IHSILTERLYVGGPMFNSKQTCGYRRCRAGAVLTTSMGNTTCYVKALAAKAGITAPT	IHSILTE	Human	2754	GenBank
2672	IRSLTERLYIGGPLTNSKGQNGYRRCRAGAVLTTSMGNTTCYKAAACRAKAKDQCT	IRSLTER	Human	2731	GenBank
2755	MLVCGDDLVVTSSEOGTEDEDERNRLAFTEATRTYSAPGDPREYVLELITSCSSNVSV	MLVCGD	Human	2814	GenBank
2732	MLVCGDDLVVTCESAGTQDEDEASLALFTEATRTYSAPGDPKKEYDELITSCSSNVSV	MLVCGD	Human	2791	GenBank
2815	ALGPGRRRYLYTRDPTPIARAMEYTRHSPVNSWLGNIIQVPTIMAMVLTTHFFSI	ALGPGR	Human	2874	GenBank
2792	AHDASGRKVVLTLDPTPLARAMEYTRHSPVNSWLGNIIMVPTIMAMVLTTHFFSI	AHDASG	Human	2851	GenBank
2875	LMADDTLDQNLNFMVGAIVSVSPDLPAITIERLHGLDASLHTYTPHELTRVASALRKL	LMADDT	Human	2934	GenBank
2852	LLAEOLEKALADCOYGCAGYSIEPLDLPQIIQRHLGSAFSLHSYSGEIRIVASCLRKL	LLAEOLE	Human	2911	GenBank
2935	GAPPLRAKSGARARVASLISRGGRRAVCGXYLFNMAVTKLKLTPLEARLLDLSMFT	GAPPLR	Human	2994	GenBank
2912	GVPLPRVRRHARARVARLLISQGRRAACGYLFLNMAVRTKLKLTPIPASQLDLSMFEV	GVPLPR	Human	2971	GenBank
2995	VGAGGGDIYHSVSRARPRLLIFGLLLFVGGLEFLPAR 3033	VGAGGG	Human	3033	GenBank
2972	AGYSGGDIYHSLSRARPRMFMWCLLLLSVGIGYILDLPR 3010	AGYSGG	Human	3010	GenBank
RESULT 14					
ABG32455					
ID ABG32455 standard; protein, 3010 AA.					
XX AC ABG32455;					
XX AC					
XX DT 15-NOV-2002 (first entry)					
XX DE Hepatitis C virus Con 1 isolate polypeptide mutant #4.					
XX KW HCV, Con 1; adaptive mutation; liver failure; cirrhosis; mutant; mutcin;					
XX KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;					
XX KW internal ribosome entry site; IRES; NS5a; HCV replication; polypeptide.					
XX OS Hepatitis C virus.					
XX OS Synthetic.					
XX FH Location/Qualifiers					

FT Misc-difference 2041 /note= "Wild-type Asn substituted by Thr"
 FT XX WO200259321-A2.
 PN 01-AUG-2002.
 PD 16-JAN-2002; 2002MO-EP000526.
 PF 23-JAN-2001; 2001US-0263479P.
 PR (RICE-) 1ST RICERCHIE BIOL MOLECOLARE ANGELETTI.
 PA De Francesco R, Magliaccio G, Paonessa G;
 PI WPI; 2002-599793/64.
 DR New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
 PT ribosome entry site (IRES) region, useful in studying HCV replication and
 PT expression.
 XX
 PS Claim 1; Page; 69pp; English.
 XX
 CC The invention relates to nucleic acid molecules comprising altered HCV
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
 CC internal ribosome entry site (IRES) region coding for one or more NS3,
 CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations
 CC are detailed in the specification. Also included are (1) an expression
 CC vector comprising a nucleotide sequence coding for the altered nucleic
 CC acid, which is transcriptionally coupled to an exogenous promoter; (2) a
 CC recombinant cell human hepatoma cell comprising the altered nucleic acids
 CC; (3) a recombinant cell produced by introducing into a human hepatoma
 CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
 CC replicon enhanced cell or which containing a functional HCV replicon; (5)
 CC an HCV replicon enhanced cells made in the method; and (6) measuring the
 CC ability of a compound to affect HCV activity. The HCV replicons and HCV
 CC replicon enhanced cells are useful in studying HCV replication and
 CC expression, and HCV and host cell interactions, producing HCV RNA and
 CC proteins, and providing a system for measuring the ability of a compound
 CC to modulate one or more HCV activities e.g. to discover drugs which may
 CC treat HCV mediated diseases such as liver failure, cirrhosis and
 CC hepatocellular carcinoma. The present sequence is the HCV replicon con 1
 CC polypeptide (comprising the Core, E1, E2, NS2, NS3, NS4A, NS4B, NS5A
 CC and NS5B proteins), NS5A mutant of the invention. Note: The present
 CC sequence is not shown in the specification but was created by the indexer
 CC using the HCV sequence appearing as ABG32451 and the information in claim
 CC 1
 CC
 XX
 SO Sequence 3010 AA;
 Query Match 74.6%; Score 12033.5; DB 5; Length 3010;
 Best Local Similarity 71.8%; Pred. No. 0;
 Matches 2183; Conservative 352; Mismatches 469; Indels 35; Gaps 7;

Db 241 ALPTLAARNASVPTTIRRHVDLVGAALCSAMVGVGLCSVFLVLAQFTFSRRHET 300
 Qy 301 VODCNCSITPGITTHGRAMDMMMMSPTATMILAYAMVPEVIIDIGAHGWNFGLA 360
 Db 301 VODCNCSITPGITTHGRAMDMMMMSPTAAVLVSQLEIRIPAQVVMVAGAHGVLAGLA 360
 Qy 361 YFSMGMAMKVVVILLLAGVDARFTHVGSAAQOTGRILTSFDMCPROKIQLVNTNGSM 420
 Db 361 YFSMGMAMKVLIVLLFAGVDGCTVYTGTAAKTIGTSLFSPSSOKIQLVNTNGSM 420
 Qy 421 HINRTALNCNDSIAHTGFIALSFYTHSFNSSGCPBRMSACRSIEAFVWGALQYEDNVTN 480
 Db 421 HINRTALNCNDSINTGFLAALFYVHKFNSSGCPBRMASCSPIDAFQGWCPITY--NES 478
 Qy 481 PEDMRPYCHNYPRQGVASATVCGPVYCFPTSPVYVCTTDRLAGPTTWGNEEDVPL 540
 Db 479 SSDORPYCMHYAPRCGIYVPAOVCPYVCFPTSPVYVCTTDRFGVPTYSWGENEDVPL 538
 Qy 541 LNSTRPPLSGFCTMNNSSGYTKTCGAPCRTRADFNASTDULCPTDCFRKHPDTTYLK 600
 Db 539 LNSTRPQGNWFGCTMNNSTGFTKTCGAPCRTRADFNASTDULCPTDCFRKHPDTTYLK 596
 Qy 601 CGSGPWLTPRCIDVYRLLMHPYCTVNYTIFKIRMYGVGBHRLTAACNFTRGDRCNLD 660
 Db 597 CGSGPWLTPRCIVHYRLLMHPYCTVNYTIFKIRMYGVGBHRLTAACNFTRGDRCNLD 656
 Qy 661 RRSQSLPLHSTTBWAILPCSYSDLPALSTGLHKNQIVYQVPMYGLSPALTKIYVM 720
 Db 657 RRSSELSPLHSTTBWQVPCSEFTLLPALSTGLHKNQIVYQVPMYGLSPALTKIYVM 716
 Qy 721 EWWILFLLLADARVACLMMLILGQAEALKEVILHAASASGNGFLYFVFAVM 780
 Db 717 EYVLLFLLLADARVACLMMLILGQAEALKEVILHAASASGNGFLYFVFAVM 776
 Qy 781 YIKGRVPLATYSLTGLMSPSLLLAPOAAVAYDASVHGOIGALLWITTEFLTPGYK 840
 Db 777 YIKGRVPLAAVAYSLTGLMSPSLLLAPOAAVAYDASVHGOIGALLWITTEFLTPGYK 836
 Qy 841 TLLSRFLMWLCULLTGLGEMVQEMAPRMQVRCGRDGIYAAVAIFPGVVFDTIKMLAVL 900
 Db 837 LFLARLIMWLQYFIRAEHLQVWIPPLNVRGRRAVILLTCAIHDELLFTYTKILLAL 896
 Qy 901 GPAYLLKGLATVPVPRVRAHALLRMCMTARHLAGRKYVMALLAGRWTFYIYDHLTM 960
 Db 897 GPAYLLKGLATVPVPRVRAHALLRMCMTARHLAGRKYVMALLAGRWTFYIYDHLTM 956
 Qy 961 SDMAASGLDLAAVEPIIFSPMEKKVIVMGAETACGDIILHGLPVSARLAGREVLGPAD 1020
 Db 957 RDMAHAGRLDLAAVEPIIFSPMEKKVIVMGAETACGDIILHGLPVSARLAGREVLGPAD 1016
 Qy 1021 GYTSKGMSLAPITAYAAQOTRGLLTIVVSMGRDKTEOAGEIQVLSYTTQSFGLSTIG 1080
 Db 1017 SLEGGQWRLAPITAYAAQOTRGLLTIVVSMGRDKTEOAGEIQVLSYTTQSFGLSTIG 1076
 Qy 1081 VLMTVYHAGNKTLLGSRGPVYOMYSABGDLVGMPSRGTSLSLEPCTCGAADLYLVNTN 1140
 Db 1077 VCMTVYHAGNKTLLGSRGPVYOMYSABGDLVGMPSRGTSLSLEPCTCGAADLYLVNTN 1136
 Qy 1141 ADVIPARRRGDKRGAALLSPRLSTLKGSSGCVLCPRGHAGVVFRAVCSRGVAKSIDPI 1200
 Db 1137 ADVIPARRRGDKRGAALLSPRLSTLKGSSGCVLCPRGHAGVVFRAVCSRGVAKSIDPI 1196
 Qy 1201 PVETLDIYTRSPFSDNSTPRAVPQTYQVGYLHAAPGSGSKTKVPAYAAQYKVLVNP 1260
 Db 1197 PVESMETTRRSVPFTDSSPPAVPQTFQVAHHAAPGSGSKTKVPAYAAQYKVLVNP 1256
 Qy 1261 SYAATLGFAGAYSKAHGIMPINRTGRTVTTGAPITYSYGEFLADGCGAGAYDIIICD 1320
 Db 1257 SYAATLGFAGAYSKAHGIDPNIRTGRTVTTGAPITYSYGEFLADGCGAGAYDIIICD 1316
 Qy 1321 ECHAVDSSTIIIGITVLDOAETAGVRLTVLATATPPGSVTTBHPNIEVALGOEIEPIFY 1380
 Db 1317 ECHSTDSSTIIIGITVLDOAETAGVRLTVLATATPPGSVTTBHPNIEVALSTGSEIPEY 1376

Qy	1381	GRALPLASTIKGRHLIFCHSKKCCDELAALARRGMLNSVAYYRGDVSVPITQODVVVA	1440
Db	1377	GKALPIETIKGRHLIFCHSKKCCDELAALSKSLGALNAVYYRGDVSVPITSDVIVVA	1436
Qy	1441	TDALMTGYTDFDSVIDCNVAVTQVVPFSLDPTFTTYTQIVPODASHSORGRTGRGL	1500
Qy	1437	TDALMTGYTDFDSVIDCNVTCQTQVTPFSLDPTFTIETTVPDVAVSORGRGTGRGM	1496
Db	1501	GIYVYVSNGERASGMFDSVLCCECDGAAAYELTPSEYTRLEAYFNTPELPLVCOHLE	1560
Qy	1497	GIYFVTPGERPSGMFDSVLCCECDGACAYELTPATSVRLBAYLNTPELPCOCHLE	1556
Db	1561	FMEAVFTGLTIDAHFLISQTKOSGENFAYLTAYOATYCARAKAPPPSDVMWKKLTBLK	1620
Db	1557	FMEVSVFTGLTIDAHFLISQTKQADGNFYLVAQATVCABAKQAPPPSDQWKKLTBLK	1616
Qy	1621	TLVGPFTPLXRLGSLVTEVTLTHPYTKIATFCQADLEWMTSTVYLAGVLAAYACLA	1680
Db	1617	TLHGPTPLXRLGAVONEVITHTHTPKYIMACMSADLEVITSTVWLVGCVLAALAAACLT	1676
Qy	1681	TGCACIIGRLHINORAVVAPDKEVLYEAFDMEBESAPPAALIEGQRALAEMLKSKIOGL	1740
Db	1677	TGSVVIYGRITLISGRPALIPREBULYRFDEMEBESHLPTIEQMOALBPQKALIGL	1736
Qy	1741	QOASKOADIQPTVQASWPKEVQFMAKHWNFISGIOYLAGLSTLPGNPAAVSMAPSA	1800
Db	1737	QTARKOAAAPAVVESKWRITLEAFMAKHMNFISGIOYLAGLSTLPGNPAAIASMAFAS	1796
Qy	1801	LTSPLSTSTTLNLNIGCMASQIAPAPGAGFVSGIYGAASIGIGKLVYLLIAGYG	1860
Db	1797	LTSPLTTOHTLFLNIGSMVAQOLAPPSAASAFGAGIAGAASIGIGKLVYLLIAGYG	1856
Qy	1861	AGISGALVAFKIMSGEKSPMEDVNNLPGIILSPGALVGVYCAAILRRHVPGEGAYQM	1920
Db	1857	AGVAGALVAFKVMGEMSTEDLVNLLPALIISPGALVGVYCAAILRRHVPGEGAYQM	1916
Qy	1921	NRLIAPASRGHNVAPTRHYTESDSQRYTOLGSLITSLRLHNMVITEDCPIPCGGSW	1980
Db	1917	NRLIAPASRGHNVAPTRHYTESDAAARTOLISLITLOLKRLHOWINEDCSPPCSGSW	1976
Qy	1981	LRDVMDVCTLLTTPKMLTSKLPFKPMGLFFVSOQKYKGVMAGTGIMTRCCGANIS	2040
Db	1977	LRDVMDMLCTVLTDFKTMLOSKLRLPBGVFFPQORIKYKGVMGDGIHOTTCGQOIT	2036
Qy	2041	GNVRLGSHRITGPYTCMNIMQCTPPINCYTEGQCVKPAENFKVAILMVAASBYAEVTH	2100
Db	2037	GHVKTGSMRIVGPRITCSNTMHGTGTPINAYTTGCPSPAPMYSAALMVAEEYEVTRV	2096
Qy	2101	GSYHYITGLTDNLIKVPQLPSPEFESVNDVOUHRPAPRKPPFRDVSVCYGLNSVY	2160
Db	2097	GDFHYVIGMTDNNKCPQVAPPEFTEVDGVRLLHRYAPACKPLIREVYTLVGLNOVLV	2156
Qy	2161	GSQLPCEPEPTDVLMSMLTDPSHITASTAARLARLARGSPPEBASASQSLASPLIRACT	2220
Db	2157	GSQLPCEBEPPVAVLTSMLTDPSHITATARRLARLARGSPPLASSASQSLASPLIKACT	2216
Qy	2221	THGKAYDVDMVDANIF----MGQDVTRIEGSGKVVYLDLDPMBEBESDLEPSIPSEYML	2276
Db	2217	TRHDSPDADLLEAMNLMQEMGNNITRVESBNKVYIIDSPEPLQAEEDBEREVSAPBILR	2276
Qy	2277	PKKRPFPALPMARPDVNPPLVESKRRDYDOPATVAGCALPPRPTTPPPRRRRYGLS	2336
Db	2277	RSKRFPFRAMPIMARPDVNPPLLESKMDDYPPVHGCPULPPAKAPPLPPRRKRYVLS	2336
Qy	2337	EDSIGDALQOALIAISFGQPPPSG-DGSLSTGAGAADSGQTPPELALSETGISISMPPL	2395
Db	2337	ESTVSSALAEIATKTGFGSESSAADSQATATA-----SPDQSDGDGADSDVESYSSMPPL	2391
Qy	2396	EGELGDDPLBEQVVEPPPPGQGVAAPOSDBSGSNSTGEF--DSDVYCCSMYSMTGALIT	2454
Db	2392	EGEGQDDPL-----SDGSMYVEEASSEDVYCCSMYSMTGALIT	2431

Qy	2455	PCSEEEKLPLNPLNSNGLRYHNKVCYCTTKRSASLRKAKYTPFMQVLDSYVSLDK	2515
Db	2432	PCAAEFTKLTINLNSNLLHNHLVYATTSRSASLRQKATYFRLQVLDDHYADVLKEMK	2491
Qy	2515	LAASKVTARLLTMEBAQCLTPPHSARSKYFGAKEVKSLSGRAVNHIKSVWKDLLEDSET	2574
Db	2492	AKASTVAKKLLSVEBAKCLTPPHSARSKFYGAKDVNLSLSSKVNHIIRSVWKDLLEBETET	2551
Qy	2575	PIPTTIWAKKEVFCVDPDTKGGKKAAARLIIVPDIGVRVCEGMAIYDIDQKLPOAVMGASYG	2634
Db	2552	PIDITTIWAKKEVFCVDPKGGKKAPARLIIVPDIGVRVCEGMAIYDVSTIPLQAVMGSSYG	2611
Qy	2635	FOYSPQARVEFLLKAAAEKDKPMGFSYDTRCFDSTVTERDIRTEESIYRACSLPEEAHTA	2694
Db	2612	FOYSPQGRVEFLVAMAKAKKCPMGFAVDTRCFDSTVTEINDIRVEESIYCCCDLAPEARQA	2671
Qy	2695	IHSITTELLVYGGPMFENSKGQTCGYRRRQASGULTTSMGNITTCYVAKLAACKAAGIAPT	2754
Db	2672	IRSLTTERLYGGPLTNSKGQNGCYRRRQASGVLITTSGNLTTCYLAKAAACRAKAKIODCT	2731
Qy	2755	MLVCGDGLVVISSESGTEEDERNLRAFTAMETYSAPGDPPEPYDLIELITSCSSNVSV	2814
Db	2732	MLVCGDGLVVISCSAGTQEDBASIRATETAMTRYSA PGDPPEPYDLIELITSCSSNVSV	2791
Qy	2815	ALGPGRRRYLLTRDPTPTIARAAMETVRS PNVSWGNITTYAPTIWMKVLTHTFFSI	2874
Db	2792	AHNASGRRVYLLTRDPTPLARAMEVARTPTPNVSWGNITIMAPTIMAMNIMTHTFFSI	2851
Qy	2875	LMADDTLDONLNFPMYGA VYVSPLDLPALIERLHGDAFSLTHTYTPHELTRVAVSALRKL	2934
Db	2852	LIMAEQLKALDCCIYACYSIEPLDLPQIILQRLHGASLSHSYSFGEINRYVASCLRKL	2911
Qy	2935	GAPPLRAWKSRARAVRASLLISRGRAAVCGRYLFENMAVTKLKLTPLEEARLLDLSWPT	2994
Db	2912	GVPEPLRVRRHARSVRARLLISGGGRATCGKYL FENMAVTRKLLKLTPIPAASQDLDSWPFV	2971
Qy	2995	VGAGGSDIYHSVSRAPRLLLFGLLLFVGVGLFLPAR	3033
Db	2972	AGYSGSDIYHSLSPARPRFMWCLLLSVGVLYLLENR	3010
RESULT 15			
ID	ABG32451	standard; protein; 3010 AA.	
XX	ABG32451;		
AC			
XX			
XX	15-NOV-2002 (first entry)		
DT			
DB			
XX			
XX	Hepatitis C virus Con 1 isolate polyprotein.		
XX			
XX	HCV; Con 1; adaptive mutation; liver failure; cirrhosis;		
KW	hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;		
KW	internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.		
KW			
OS	Hepatitis C virus.		
XX			
XX	MO200259321-A2.		
XX			
XX	01-AUG-2002.		
PD			
XX	16-JAN-2002; 2002MO-EP000526.		
XX			
XX	23-JAN-2001; 2001US-0263479P.		
XX			
PA	(RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.		
XX			
PI	De Francesco R, Migliaccio G, Paonessa G;		
XX			
DR	WPI: 2002-599793/64.		
XX			
DR	N-PSDB; ABK91411.		
XX			
PT	New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV		

PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication and
XX expression.
XX Claim 1, Page 34-36; 69pp; English.
CC The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations
CC are detailed in the specification. Also included are (1) an expression
CC vector comprising a nucleotide sequence coding for the altered nucleic
CC acid, which is transcriptionally coupled to an exogenous promoter; (2) a
CC recombinant cell human hepatoma cell comprising the altered nucleic acids
CC : (3) a recombinant cell produced by introducing into a human hepatoma
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
CC replicon enhanced cell or which containing a functional HCV replicon; (5)
CC an HCV replicon enhanced cells made in the method; and (6) measuring the
CC ability of a compound to affect HCV activity. The HCV replicons and HCV
CC replicon enhanced cells are useful in studying HCV replication and
CC expression, and HCV and host cell interactions, producing HCV RNA and
CC proteins, and providing a system for measuring the ability of a compound
CC to modulate one or more HCV activities e.g. to discover drugs which may
CC treat HCV mediated diseases such as liver failure, cirrhosis and
CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
CC polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A
CC and NS5B proteins) used as a basis for the adaptive mutations of the
XX invention
XX

SQ Sequence 3010 AA;

Query Match 74.6%; Score 12031.5; DB 5; Length 3010;
Beet Local Similarity 71.8%; Pred. No. 0;
Matches 2183; Conservative 352; Mismatches 469; Indels 35; Gaps 7;

QY 1 MSTNPKQKRTKNTNRPRQDVKFGGQIVGVVLLPRRPRGLGRATRKTSERSQPRG 60
DB 1 MSTNPKQKRTKNTNRPRQDVKFGGQIVGVVLLPRRPRGLGRATRKTSERSQPRG 60
QY 61 RROPIPKDRSTGKSGKPGYWPPLYGNBGLGAGWLLSPRSGRPSGWNDRPHSRNNG 120
DB 61 RROPIPKARQPEGRARAAQPGYWPPLYGNBGLGAGWLLSPRSGRPSGWPDPDRRRSRNNG 120
QY 121 KYIDLTGFGFADLMGYIPVVGAPLGGVARALAHGVVLEDDGVNFAAGNLPGCSFSIFLLA 180
DB 121 KYIDLTGFGFADLMGYIPVVGAPLGGVARALAHGVVLEDDGVNFAAGNLPGCSFSIFLLA 180
QY 181 LLSCTITPVSAAEVKNISTGYMTNDCTNDSITWQLOAAVLHVPGCVPEKYGNAASQCVI 240
DB 181 LLSCTITPVSAAEVKNISTGYMTNDCTNDSITWQLOAAVLHVPGCVPEKYGNAASQCVI 240
QY 241 PVSNNVAVORPGALFGRLTHIDMVVMSATLCSALVYGDLCGVMILAAQMTVSPQHMF 300
DB 241 ALTPITLAANASVPTTITRRHYDLVGAALCSAMVYGDLCGSVFVLAQLFTFSRPHRET 300
QY 301 VDDCNCSITPRTITGRRAMDMMMNNSPTATMILAAARVPEVIIDISGAMGWRGLA 360
DB 301 VDDCNCSITPRTITGRRAMDMMMNNSPTATMILAAARVPEVIIDISGAMGWRGLA 360
QY 361 YFSMOGAAKAVVLLILLAGVDAARTHTVGSAAQTGLTSLFDMQPRKIQLVNTNGSM 420
DB 361 YFSMOGAAKAVVLLILLAGVDAARTHTVGSAAQTGLTSLFDMQPRKIQLVNTNGSM 420
QY 421 HINRRLANCNDLSLHTGFIASLFYTHSFGSGCEBMSACRSIBAFVNGALQYEDNVTN 480
DB 421 HINRRLANCNDLSLHTGFIASLFYTHSFGSGCEBMSACRSIBAFVNGALQYEDNVTN 480
QY 481 PEDMRPYCHYPRGCGVSAKTVGSPVVCFTPSPVVGTGTRDLGAPYTWKBNETDVL 540
DB 481 PEDMRPYCHYPRGCGVSAKTVGSPVVCFTPSPVVGTGTRDLGAPYTWKBNETDVL 540
QY 541 LNSTRPGLGSWFGCTMMNSGYTKTCGAPPCRTADFNASTDLCTPDCFRRKHPDTTYLK 600
DB 541 LNSTRPGLGSWFGCTMMNSGYTKTCGAPPCRTADFNASTDLCTPDCFRRKHPDTTYLK 600

DB 539 LNSTRPQGNWFGCTMMNSGTGTCGPPCNIGIGNKT--LTCPDCCFRKHPATYTK 596
QY 601 CGSGFWLTPRCILIDVRYLMMHYPCVTWNTIEXIRBMVGVGVEHRLTAACNPTGDCNLED 660
DB 597 CGSGFWLTPRCILVHYRMLMHPCTVNFITFVRYVGVGVEHRLTAACNPTGDCNLED 656
QY 661 RDRSOLPILNHTSTTMALIPCSYSDLPALSTGLLHQIVVQFMVYGSPLATKYIVM 720
DB 657 RDRSELSPILSTTEMQVLPSCFTLLPALSTGLILHQVNVQVLYGIGSAVNSPAIKM 716
QY 721 EVVILLPLILADARVACILMMLILGOBALEKLVIIHAASASCNGLYVVFVNA 780
DB 717 EVVILLPLILADARVACILMMLILGOBALENIIVINASVAAHGISLTFECCAM 776
QY 781 YKGVVPLATSLGLMSFSLILALQOAAVYASVHGOIQAALLVMTLPTLRGK 840
DB 777 YKGVVPLAAVALYGVWPELILLLALPPRAAMREMAASCGAVFGLILTLTSPHK 836
QY 841 TLLSFLMWLCYLLTYGEMVQENAPPMQVRGGRDGIWAVAIYPGVVFDTIKMLAVL 900
DB 837 LPLALLIMWLYFTIRABNHLQWILPILNVGGRVAILLTCIAHPELITTKILLAL 896
QY 901 GPAYLLKGLATREVPYFVRAHALRMCTWARHLAGRYVQMALALGRWTGYIYDHLTPM 960
DB 897 GPLMWLQGITKVPYFVRAHGLIRACMLVRAKVAAGHYVQMALMKLALGTIVVDHLPL 956
QY 961 SPMAASGLDLAVANPEIIFSEMEKKVYWGMEFAACGDILHGLVYSARLGEVLLGPD 1020
DB 957 RDMAHAGRLDLAVANPEVVFSDMETKVTWGDATACDIIILGLVPSARGEHLGPD 1016
QY 1021 GYTSKWSLAPITTYAQQTRGLGTIVSMTRKDEQAGEIIOVLTYSQSFGLSTSSG 1080
DB 1017 SLEGGWMLAPITTYAQQTRGLGTITSLGRNRNVEGVEVQVSTATGTFLATCVNG 1076
QY 1081 VLMTVYHAGNKTLAGSRQPVTOYMSABGDIWGPSPGTSLEPCTCGADVLYLVN 1140
DB 1077 VCMTVYHAGSKTLGAPKPIOTOMYTNVDDQVGAQPPARSLPTCTGSSDLYLVTRH 1136
QY 1141 ADVIFARRRGDRGALLSPRLSTLKSSGGGVLCPRHNAVCPRAACSRVAASIDPT 1200
DB 1137 ADVIFARRRGDRGALLSPRLSTLKSSGGGVLCPRHNAVCPRAACSRVAASIDPT 1196
QY 1201 PVETLADIYTRSPFTDNSTTPPAVPTQYVYLAHATGSGKSTKVVAAYAAQGYKVLVLP 1260
DB 1197 PVESMETMRSVFTDNSSPAVPOTFOVAHHAATGSGKSTKVVAAYAAQGYKVLVLP 1256
QY 1261 SVAATLGFQAYLSKAHGINPNIRTVTTGAPITTYSTYKFLADGGCAGAYDIICD 1320
DB 1257 SVAATLGFQAYVMSKAHIDPNIRTVTTGAPITTYSTYKFLADGGCAGAYDIICD 1316
QY 1321 ECHAVDSTTIIGITGVLDQAEAGVRLVLTATTPGSVTTHPHNIEVYALQOBEIIPY 1380
DB 1317 ECHAVDSTTIIGITGVLDQAEAGVRLVLTATTPGSVTTHPHNIEVYALQOBEIIPY 1376
QY 1381 GRAIPLSYIKGGRHLIFCHSKKKKDELAALARGMLNSAAYRGADVSIPTQGVVYA 1440
DB 1377 GRAIPLEITIKGGRHLIFCHSKKKKDELAALARGMLNSAAYRGADVSIPTQGVVYA 1436
QY 1441 TDALMTGYTGFDSYIDCNVAVTQVVDSPDLPTFTITTOIVQDVAVSSORRGRTGRGL 1500
DB 1437 TDALMTGFGDPDSYIDCNVAVTQVVDSPDLPTFTITTOIVQDVAVSSORRGRTGRGL 1496
QY 1501 GYIRYVTSGERASGNFDSVYLCECDAGAAWELTPSETTYRLAAYPTPGLPYQODHLE 1560
DB 1497 GYIRYVTPRERSGNFDSVYLCECDAGAAWELTPSETTYRLAAYPTPGLPYQODHLE 1556
QY 1561 FWEAVFTGLTHIDAFILSOTKOSGENFAYLTAYQATVABAKAPPSMDVMKCLTRLKP 1620
DB 1557 FWEAVFTGLTHIDAFILSOTKOSGENFAYLTAYQATVABAKAPPSMDVMKCLTRLKP 1616
QY 1621 TLVGPPLLYRIGSVTNEVTLTHPVTKYIATCMQADLEVMSTVWLAGVLAAYACLA 1680
DB 1617 TLHGPPLLYRIGAVQNEVTTTHPTTKYIMACMSADLEVMSTVWLAGVLAAYACLA 1676

QY 1661 TGCVCITIGRLHINORAVAPDXEVLVEAFDEMECCASRAALIEGQRIANMLKXIOGLL 1740
Db 1677 TGSVVIYVGRILISGRPAIIPDREVLVYREFDEMECCASHLPYIEQOMQOLAEQKOKAIGLL 1736
QY 1741 QOASQAOADIOPTVOASMPKVQFOFAKXHMNFISGIOYAGISTLPGNAVAVSMAFSA 1800
Db 1777 QTAHQAEAAAFVVSXKRTLEAFMAKHMNFISGIOYAGISTLPGNALASLAFAS 1796
QY 1801 LTPSLSTSTTLLNLLIGWLAQIAPPAAGATFVVSGLVGAAGVSGIGKVLVDLLAGY 1860
Db 1797 ITSPITLQHTLLFNLLIGWVAQOLAPPSAASAFVGAAGIAGAAGVSGIGKVLVDLLAGY 1856
QY 1861 AGISGALVAFKIMSGEKPSMEDVNNILPGIISPGALVGVICAIILRRHVGEBAVQNM 1920
Db 1857 AGVAGALVAFKVMGSEMPSTEDLVNLLPALISPGALVGVVCAALIRRHVGEBAVQNM 1916
QY 1921 NRIIPASRGNHVAAPHVYTESDASQRYTQOLGSLITISLLRLNMTEDCPICGGM 1980
Db 1917 NRIIPASRGNHVSPTHVYTESDAAARVQILSLITQILRLHQMINECSTPCSGW 1976
QY 1981 LRDVADWVCTILTFKXNMLTSKLPKMPGLPFVSCQKGYGVMAGTIGMTTRCPGAGNIS 2040
Db 1977 LRDVADWVCTVLTDFKXNMLTSKLPKMPGLPFVSCQKGYGVMAGTIGMTTRCPGAGNIS 2036
QY 2041 GNVRLGSMRITGPKTCOMINMOCTPINCYTEGOCVPKPAPNFKVALMRYAASEYAVTQH 2100
Db 2037 GHVKNKSMRIVGPRICSNMTMGTFPINAYTTGCTPSPAFNYSRALMRYAASEYAVTQV 2096
QY 2101 GSYHYITGILTNMLKVPOLSPPEFSWQVQIHRFAPTKPFRDEVSFCVGLNSFVY 2160
Db 2097 GDFHYVTGTDNVKPCQVPAPEFTBEDVGLHRYAPACKPRLREBVTFLVGLNQYLV 2156
QY 2161 GSQLPCEPDPDVLMSMLTDPSSHITAEYARLARSPSEASSASQLSAPSLRATCT 2220
Db 2157 GSQLPCEPDPDVAVLTSMLTDPSSHITAEYARLARSPSEASSASQLSAPSLRATCT 2216
QY 2221 THGKAYVDVMDANLF---MGQDVTRIISGSKRVVLDLDPMEBERSDLEPSISEYML 2276
Db 2217 THHDSPPDADLIEBANILMROBMGNITRVSENNKVILDSFEPLQAEDEDEREVSVAELIR 2276
QY 2277 PKRPPPALPAMARPDYNNPLVESWKRPDYOPATVAGCALPPRKTPTPPRRRTVGLS 2336
Db 2277 PKRPPPALPAMARPDYNNPLVESWKRPDYOPATVAGCALPPRKTPTPPRRRTVGLS 2336
QY 2337 EDSIGDALQOLAIKSGQPPSG-DGSLGTGAGAADSQTPPDELALSETGSISSMPL 2395
Db 2337 ESTVSSALAEALTKTTGSSBSAIVDSGTATA-----SPDQPSDDDGAGSDVESYSMPL 2391
QY 2396 EGEIGDPDLEPQVEBPQPPQGVAAFGSDSGSWSTCSEE-DDSVVCCSMYSWTGALIT 2454
Db 2392 EGEIPDPLD-----SDSWSWTSREASBDVCCSMYSWTGALIT 2431
QY 2455 PCSPEBEKLPINPLNSILRYNNKYCTTKSASLRAKVTFRDMOVLDSYYDSVLKDIK 2514
Db 2432 PCAABETKLPINALSNSILRHNLVYATTSRASLSQKKVTFDRLQVLDHYRDVLKEMK 2491
QY 2515 LAASKVTARLLTMEACQLTTPHSASRKYFGAKKEYRSLSGRAVNHKISVMKDLLEDSBT 2574
Db 2492 AKASTYKATLSVEBECKLTTPHSASRKYFGAKKEYRSLSGRAVNHKISVMKDLLEDTBT 2551
QY 2575 PIPTTITMANEVEFCVDPFKGKKAARLIYVYVPCVCEKMAIYDITQKLPQAVMGASYG 2634
Db 2552 PIDTITMANEVEFCVQPEKGRKPARLIYVYVPCVCEKMAIYDVSTLPQAVMGASYG 2611
QY 2635 FOYSPAQRVTEFLKAAEKDPMGFSYDRCPESTYTERDITREBSIYACSLPEBAHTA 2694
Db 2612 FOYSPAQRVTEFLKAAEKDPMGFSYDRCPESTYTERDITREBSIYACSLPEBAHTA 2671
QY 2695 IHSILTERLYVGGPMFNSKQOTCGYRRCRASGVLTSMGNTITCYVVALAACKAAGIIAPT 2754
Db 2672 IHSILTERLYVGGPILNSKQNGCYRRCRASGVLTSMGNTITCYVVALAACKAAGIIAPT 2731

QY 2755 MLVCGDDILVVISSEQTEBERNLRAPTEAMTRYSAAPGDPPEPPEYDLELITGSSNVSV 2814
Db 2732 MLVCGDDILVVISSEAGTQEDBASLRAFTAMTRYSAAPGDPPEPPEYDLELITGSSNVSV 2791
QY 2815 ALGPGRRRYLTLDPTPIARAAMETVRHSPVNSWLNIIQYAPTIMARVNLMTHPFSI 2874
Db 2792 AHDASGKRYLTLDPTPIARAAMETVRHSPVNSWLNIIQYAPTIMARVNLMTHPFSI 2851
QY 2875 LMAQDTLDQNLFEWYGAVVSPLDPAIIBRLHGLDAFSIHTYTPHBLTRVASALRKL 2934
Db 2852 LLAQDEKALDCQYGAACYSIBPLDPOIIRLHGLSAFSLHSYSGEINRVASCLRKL 2911
QY 2935 GAPPLRAKRSARAVASLIRSGRAAVCSKYLPMWAVKTLKLTPLPEARLLDSSWFT 2994
Db 2912 GVPPLRVHRHARSYRARLLSQGGRAATCGKXLFPMWAVRTKLKLTPIPAASQDLSSWFT 2971
QY 2995 VAGAGGDIYHSVSRARPRLLFGLLLPVGYGLFLLPAR 3033
Db 2972 AGYSGGDIYHSLSRARPRMFMCLLLSVGVGYLLPNR 3010

Search completed: October 28, 2005, 15:59:04
Job time : 304 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2005, 15:48:07 ; Search time 75 Seconds

(without alignments)
3891.006 Million cell updates/sec

Title: US-09-980-559-2

Perfect score: 16120

Sequence: 1 MSTNPKPQKRTKRNTRRPQ.....LFGILLFVGIVGLFLPAR 3033

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : PIR_79:*

1: pirl:*

2: pirl:*

3: pirl:*

4: pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15830	98.2	3033	1 JQ1303	genome polypeptide
2	13971	86.7	3033	1 GNMVU8	genome polypeptide
3	11998.5	74.4	3010	1 GNMVU3	genome polypeptide
4	11984.5	74.3	3010	1 GNMVU3	genome polypeptide
5	11979.5	74.3	3010	1 A45573	genome polypeptide
6	11922	74.0	3011	1 GNMVU3	genome polypeptide
7	11911.5	73.9	3010	1 GNMVU3	genome polypeptide
8	11885	73.8	3011	1 S40770	genome polypeptide
9	11843.5	73.5	3010	1 S18030	genome polypeptide
10	11796	73.2	3011	1 GNMVU3	genome polypeptide
11	11500.5	71.3	3014	1 JCS681	genome polypeptide
12	3490.5	21.7	874	2 J00881	genome polypeptide
13	3263	20.2	876	2 PC2319	polypeptide - hepa
14	3231	20.0	872	2 S18875	genome polypeptide
15	3214	19.9	782	2 S18876	genome polypeptide
16	3194	19.8	782	2 S18032	genome polypeptide
17	3194	19.8	782	2 S18031	genome polypeptide
18	3126	19.4	787	2 PN0677	hypothetical prote
19	2937.5	18.2	874	2 J00883	genome polypeptide
20	2671	16.6	640	2 JQ1584	genome polypeptide
21	2596	16.1	716	2 JQ1366	polypeptide - hepa
22	2199	13.6	3005	2 T08841	polypeptide - hepa
23	2114.5	13.1	2970	2 T08839	polypeptide - hepa
24	2111.5	13.1	520	2 JQ1925	polypeptide - hepa
25	2108.5	13.1	523	2 JQ1926	polypeptide - hepa
26	2099	13.0	513	2 PC1284	genome polypeptide
27	2089	13.0	550	2 JH0711	genome polypeptide
28	2074	12.9	492	2 PS0326	polypeptide - hepa
29	2056	12.8	513	2 A44150	structural protein

30	1958.5	12.1	492	2	S41288	genome polypeptide
31	1808	11.2	441	2	S12707	genome polypeptide
32	1761	10.9	350	2	S35631	genome polypeptide
33	1734.5	10.8	386	2	S68016	ATPase/RNA helicas
34	1630	10.1	411	2	PC2061	genome polypeptide
35	1622	10.1	411	2	PC2060	genome polypeptide
36	1486.5	9.2	369	2	S21471	genome polypeptide
37	1451	9.0	415	2	PC4407	envelope protein -
38	1450	9.0	365	2	J00880	NS5 protein - hepa
39	1430	8.9	365	2	J00879	NS5 protein - hepa
40	1376	8.5	322	2	JN0265	genome polypeptide
41	1085	6.7	315	2	PS0165	envelope glycoprot
42	1076	6.7	315	2	PN0011	envelope glycoprot
43	1062	6.6	315	2	PS0164	envelope glycoprot
44	984	6.1	216	2	S21337	genome polypeptide
45	962	6.0	234	2	S32742	genome polypeptide

ALIGNMENTS

RESULT 1

JQ1303 genome polypeptide - hepatitis C virus (isolate HC-J6)
N:Contains: capsid protein C; envelope protein M; hepatitisin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: JQ1303
R:Okamoto, H.; Okada, S.; Sugiyama, Y.; Kurai, K.; Iizuka, H.; Machida, A.; Miyakawa, Y.
J. Gen. Virol. 72, 2697-2704, 1991
A:Title: Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a hum
A:Reference number: JQ1303; MUID:920444440; PMID:1658196
A:Accession: JQ1303
A:Molecule type: genomic RNA
A:Residues: 1-3033 <OKA>
A:Cross-references: UNIPROT:P26660; GB:D00944; NID:9221650; PIDN:BAA00792.1; PID:9221651
A:Experimental source: isolate HC-J6 from a Japanese individual
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: ATP; glycoprotein; hydrolyase; P-loop; polypeptide; serine proteinase; transmem
F:2-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EBM>
F:192-389/Product: major envelope protein E #status predicted <NS1>
F:390-733/Product: nonstructural protein NS1 #status predicted <NS1>
F:734-1010/Product: nonstructural protein NS2 #status predicted <NS2>
F:1011-1619/Product: nonstructural protein NS3 #status predicted <NS3>
F:1316-1321/Region: DEXH motif
F:1320-1323/Region: DEXH motif
F:1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>
F:1867-2017/Product: nonstructural protein NS4b #status predicted <N4B>
F:2018-3033/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,224,305,325,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,2038,28

Query Match

Best Local Similarity 98.2%; Score 15830; DB 1; Length 3033;
Matches 2967; Conservative 30; Mismatches 36; Indels 0; Gaps 0;

QY	1	MSTNPKPQKRTKRNTRRPQDYKFGGQIVGVLLRRGRGLGVRAITRKSERSOPRG	60
DB	1	MSTNPKPQKRTKRNTRRPQDYKFGGQIVGVLLRRGRGLGVRAITRKSERSOPRG	60
QY	61	PROPIPKDORSTGKSGKRGYPWPLVYGNEGGLMAGMLSPQSRPSGSGNDPRHRSRVNG	120
DB	61	RROPIPKDORSTGKSGKRGYPWPLVYGNEGGLMAGMLSPQSRPSGSGNDPRHRSRVNG	120
QY	121	KVIDLTGCGFADLMGYIPVGGAPLGGVARALAHGVRVLEDGVNFAATGNLPGCSFSIFLLA	180
DB	121	KVIDLTGCGFADLMGYIPVGGAPLGGVARALAHGVRVLEDGVNFAATGNLPGCSFSIFLLA	180
QY	181	LLSCTTVVSAAEVNISTGVMTNDCTNDSITWQLAAVLAHVPCVCCEKGNASQCM	240
DB	181	LLSCTTVVSAAEVNISTGVMTNDCTNDSITWQLAAVLAHVPCVCCEKGNASQCM	240

QY 241 PVS PNVAVOPGALTOGLRTHIDMVMSATLCSALYVGDLCGGVMLAQMFIYSPQHMF 300
 Db 241 PVS PNVAVOPGALTOGLRTHIDMVMSATLCSALYVGDLCGGVMLAQMFIYSPQHMF 300
 QY 301 VQCNCSYIPGTTTGRHMAWMMNWSPTATMLTAAVRPEYITDIIGAHKGVWFGLA 360
 Db 301 VQCNCSYIPGTTTGRHMAWMMNWSPTATMLTAAVRPEYITDIIGAHKGVWFGLA 360
 QY 361 YFSMOMAKAVVYLLLAAGVDARTHTVGSSAAQTGRLTSLFDMQPROKIQLVNTNGSM 420
 Db 361 YFSMOMAKAVVYLLLAAGVDARTHTVGSSAAQTGRLTSLFDMQPROKIQLVNTNGSM 420
 QY 421 HIRKTLNCNDSJLHTGFASLFTTHSPNSGCCPERMSACRSIPAFVWGALQYEDNVTN 480
 Db 421 HIRKTLNCNDSJLHTGFASLFTTHSPNSGCCPERMSACRSIPAFVWGALQYEDNVTN 480
 QY 481 PEDMRPYCMHYPPOGCVSAKTVCGPVYCFTPSPVAVGTBRLGAPTYTWGENEDVFL 540
 Db 481 PEDMRPYCMHYPPOGCVSAKTVCGPVYCFTPSPVAVGTBRLGAPTYTWGENEDVFL 540
 QY 541 LNSTRPPLGSMFCCTMNNSSGYTKCGAPCPTRADFNASTDULCPTDFRKHDPDTYLK 600
 Db 541 LNSTRPPLGSMFCCTMNNSSGYTKCGAPCPTRADFNASTDULCPTDFRKHDPDTYLK 600
 QY 601 CGSGPMLTPRCLIDYPRMLMHPCTVANNYTI FKIRMTVGVYEHRLTAACNFTGRDCNLED 660
 Db 601 CGSGPMLTPRCLIDYPRMLMHPCTVANNYTI FKIRMTVGVYEHRLTAACNFTGRDCNLED 660
 QY 661 RDSQSPLSHSTTEWAILPCSSDLPAISTGILHONVVOVFMYGSLPATKTIYVM 720
 Db 661 RDSQSPLSHSTTEWAILPCSSDLPAISTGILHONVVOVFMYGSLPATKTIYVM 720
 QY 721 EWWYLLFLLLADARVACLMMLILLOAEALAEKLYILHAASNASNGELYEYIFVFAAM 780
 Db 721 EWWYLLFLLLADARVACLMMLILLOAEALAEKLYILHAASNASNGELYEYIFVFAAM 780
 QY 781 YIKGRVPLATYSLTGLMSFSLLLALPOQAVAYDASVHQIGALLVMTLETTLPYK 840
 Db 781 YIKGRVPLATYSLTGLMSFSLLLALPOQAVAYDASVHQIGALLVMTLETTLPYK 840
 QY 841 TLISRLMMLCYLLTLEAMVOMAPPMOVRGGRDGIIMAVALFYCGVFDITKMLLAVL 900
 Db 841 TLISRLMMLCYLLTLEAMVOMAPPMOVRGGRDGIIMAVALFYCGVFDITKMLLAVL 900
 QY 901 GPAYLLKAGALTVPYFVRAHALRMCTMAHLAGRYVQWALLALGRMTGYIYDLITPM 960
 Db 901 GPAYLLKAGALTVPYFVRAHALRMCTMAHLAGRYVQWALLALGRMTGYIYDLITPM 960
 QY 961 SDMAASGLRDLAVAVEPIIFSPMEKXYIWGAEFTAACGDIHLGLPVSARLGREVLLGPD 1020
 Db 961 SDMAASGLRDLAVAVEPIIFSPMEKXYIWGAEFTAACGDIHLGLPVSARLGREVLLGPD 1020
 QY 1021 GYTSKMSLAPITAYAOQTRGLTIVSMTGRDTEOAGEIQVISTYQSLTGLTISG 1080
 Db 1021 GYTSKMSLAPITAYAOQTRGLTIVSMTGRDTEOAGEIQVISTYQSLTGLTISG 1080
 QY 1081 VLTWVHGAGNKTLAGSRGPVTOMYSABEDLVGWSPTKSLFECTCGADVLYVTN 1140
 Db 1081 VLTWVHGAGNKTLAGSRGPVTOMYSABEDLVGWSPTKSLFECTCGADVLYVTN 1140
 QY 1141 ADVIIPARRRDKXGALLSPRLSTLKSSGGPVLCPRGHAGVFRAAVCSRGVAKSIDPT 1200
 Db 1141 ADVIIPARRRDKXGALLSPRLSTLKSSGGPVLCPRGHAGVFRAAVCSRGVAKSIDPT 1200
 QY 1201 PVETLIDVTRSPTFSDNSTPAVPQTYOVGYLHAPGSGKSTVPVAVYAAOGKVLVMP 1260
 Db 1201 PVETLIDVTRSPTFSDNSTPAVPQTYOVGYLHAPGSGKSTVPVAVYAAOGKVLVMP 1260
 QY 1261 SVATLIGFAYLKAHGININIRTVTGGAPITYSTGKFLADGCGAGAYDIIICD 1320
 Db 1261 SVATLIGFAYLKAHGININIRTVTGGAPITYSTGKFLADGCGAGAYDIIICD 1320
 QY 1321 ECHAVDSTTILIGITVLDQAEFTAGVRLTVLATATPPGCVTTPHPNIEVALGQGEHIFPY 1380

Db 1321 ECHAVDSTTILIGITVLDQAEFTAGVRLTVLATATPPGCVTTPHPNIEVALGQGEHIFPY 1380
 QY 1381 GRATPLSTIKGRHLIFCHSKKKCDELLAALRGGLNSVAAYRGLDVSIFPTQDDVVVA 1440
 Db 1381 GRATPLSTIKGRHLIFCHSKKKCDELLAALRGGLNSVAAYRGLDVSIFPTQDDVVVA 1440
 QY 1441 TDALMTGYGDPDSVIDCNVAVTQVVPFSLDPTFTITTOIVPODASRSQRRGTGRGL 1500
 Db 1441 TDALMTGYGDPDSVIDCNVAVTQVVPFSLDPTFTITTOIVPODASRSQRRGTGRGL 1500
 QY 1501 GIYRYVSTGERASGMPFDSVULCECYDAGAAHYELTSPSTYRLAAFPYTPCLPYCOHLE 1560
 Db 1501 GIYRYVSTGERASGMPFDSVULCECYDAGAAHYELTSPSTYRLAAFPYTPCLPYCOHLE 1560
 QY 1561 FMEAVFTGLTHIDHAFISQTKOSGNFAYLTAAYOATVCARAKAPPSMDVMKCLTRLRP 1620
 Db 1561 FMEAVFTGLTHIDHAFISQTKOSGNFAYLTAAYOATVCARAKAPPSMDVMKCLTRLRP 1620
 QY 1621 TLVGPPLLYRLGSVTYNEVLTJHPYTKXIATCMQADLEVMSTWVLAGVLAAYVCLA 1680
 Db 1621 TLVGPPLLYRLGSVTYNEVLTJHPYTKXIATCMQADLEVMSTWVLAGVLAAYVCLA 1680
 QY 1681 TGCVCITIGRLHINORAVVAPDKXYLYEAFDEMECASPAALIEBCORIABLKSKIQGL 1740
 Db 1681 TGCVCITIGRLHINORAVVAPDKXYLYEAFDEMECASPAALIEBCORIABLKSKIQGL 1740
 QY 1741 QOASKOADDIPTYOASMPKIEOPFAKMMNFISGIQYLAQSLTLPGNPAVAMMASPA 1800
 Db 1741 QOASKOADDIPTYOASMPKIEOPFAKMMNFISGIQYLAQSLTLPGNPAVAMMASPA 1800
 QY 1801 LTPSLSTTILILNLGMLASQIAPAGATGVVSGLVGAAGVSGIGKVLVDIILGYG 1860
 Db 1801 LTPSLSTTILILNLGMLASQIAPAGATGVVSGLVGAAGVSGIGKVLVDIILGYG 1860
 QY 1861 AGISGALVAFKINGEKSMEDEVNLLPGLISPGALVGVCAAILLRHVPGGCAVQM 1920
 Db 1861 AGISGALVAFKINGEKSMEDEVNLLPGLISPGALVGVCAAILLRHVPGGCAVQM 1920
 QY 1921 NRLIAPASGNHVAPTHVTESDASORVTOLGSLTITSLRLHNNITTEDCPIPCGSM 1980
 Db 1921 NRLIAPASGNHVAPTHVTESDASORVTOLGSLTITSLRLHNNITTEDCPIPCGSM 1980
 QY 1981 LRDVMDWCTITLTFKMLTJSKLPFKMPGLPFVSCQKYGKVMAGTGMTTRCPGANIS 2040
 Db 1981 LRDVMDWCTITLTFKMLTJSKLPFKMPGLPFVSCQKYGKVMAGTGMTTRCPGANIS 2040
 QY 2041 GNVRLGSKRITGPRTCKNIWGTPEPINCYTGGQCVKPAKPAKIALMRVAASEAVTQH 2100
 Db 2041 GNVRLGSKRITGPRTCKNIWGTPEPINCYTGGQCVKPAKPAKIALMRVAASEAVTQH 2100
 QY 2101 GSYHYITGLTJTDNLKVPQOLSPPEFSSWVDVQIHRFAPTRKPPFRDVSFCVGLNSFV 2160
 Db 2101 GSYHYITGLTJTDNLKVPQOLSPPEFSSWVDVQIHRFAPTRKPPFRDVSFCVGLNSFV 2160
 QY 2161 GSQJPCDEPDPDVLTSMLTDPHSHTAETAARLARSGSPSEASSASQJAPSLRATCT 2220
 Db 2161 GSQJPCDEPDPDVLTSMLTDPHSHTAETAARLARSGSPSEASSASQJAPSLRATCT 2220
 QY 2221 THGKAYDVDMVDANI FMSGDVTRIESGSKVYVLDLSPMEERSDLEPSISEYMLPKKR 2280
 Db 2221 THGKAYDVDMVDANI FMSGDVTRIESGSKVYVLDLSPMEERSDLEPSISEYMLPKKR 2280
 QY 2281 PPPALPAMARPDVNPPELVESMKRPDYOPATYAGCALPPPKRTPPPPRRRTTGVLSSESI 2340
 Db 2281 PPPALPAMARPDVNPPELVESMKRPDYOPATYAGCALPPPKRTPPPPRRRTTGVLSSESI 2340
 QY 2341 GDALQQLAIKSFQGPSPSGDGLSTGAGANDSGSTPPDELALSETGSISSMPLLEGRLG 2400
 Db 2341 ADALQQLAIKSFQGPSPSGDGLSTGAGANDSGSTPPDELALSETGSISSMPLLEGRLG 2400
 QY 2401 DPDLBPVBEPOPPOGCVAAAPGSDSGSWTCSBEDSDVVCSSMSYMTGALLTPCSBE 2460

Db 2401 DDELEPEQVLEPPPGQGVVTPGSGSGSSTCEEDDSVVCSSMSYMTGALITPCSPRE 2460
QY 2461 EKLPIINPLNSILRRYHNKYCTTTKSASIRAKKVFEDRMQVLDSDYDVLKDIKLAASRY 2520
Db 2461 EKLPIINPLNSILRRYHNKYCTTTKSASIRAKKVFEDRMQVLDSDYDVLKDIKLAASRY 2520
QY 2521 TARLLTMEBAQCLTPPHSARSRYGCAKEVRSLSGRAVNHISVWMDLDESEPIPTTI 2580
Db 2521 TARLLTMEBAQCLTPPHSARSRYGCAKEVRSLSGRAVNHISVWMDLDESEPIPTTI 2580
QY 2581 MAKNEVFCDPTPKGKGKARLIYVPDLGYRVCERKALYITQKLPQAVKASGYFOYSPA 2640
Db 2581 MAKNEVFCDPTPKGKGKARLIYVPDLGYRVCERKALYITQKLPQAVKASGYFOYSPA 2640
QY 2641 QVVEFLTKMAKADKPMGFSYDTRCDSVTTERDITRTEISYRACSLPEEATAHSLTE 2700
Db 2641 QVVEFLTKMAKADKPMGFSYDTRCDSVTTERDITRTEISYRACSLPEEATAHSLTE 2700
QY 2701 RLTVGGPMFNSKQOTCGYRCRASGVLTTSMGNTTTCYKALAAKAGIIPATMLVCGD 2760
Db 2701 RLTVGGPMFNSKQOTCGYRCRASGVLTTSMGNTTTCYKALAAKAGIIPATMLVCGD 2760
QY 2761 DLVISESGTEDEERNILAFTEAMTRYSAAPGDPPEPYDELITSCSSNVVALGPGQ 2820
Db 2761 DLVISESGTEDEERNILAFTEAMTRYSAAPGDPPEPYDELITSCSSNVVALGPGQ 2820
QY 2821 RRRYLTDPPTPIRAANETVRSVNSWLNIIQYATIWAWMLTHFSSILMAQOT 2880
Db 2821 RRRYLTDPPTPIRAANETVRSVNSWLNIIQYATIWAWMLTHFSSILMAQOT 2880
QY 2881 LDQNFEMFYGAVSVSPDLPAIETRLHGLDAFSLHTYTPHELTFSVASALRKLGPPLR 2940
Db 2881 LDQNFEMFYGAVSVSPDLPAIETRLHGLDAFSLHTYTPHELTFSVASALRKLGPPLR 2940
QY 2941 AKKSPRAVAVASLISRGGAACVGRYLFNMAVVKTKLTPLEBARLLDSSWFTVAGGG 3000
Db 2941 AKKSPRAVAVASLISRGGAACVGRYLFNMAVVKTKLTPLEBARLLDSSWFTVAGGG 3000
QY 3001 DIYHSVRARPRLLFGILLTFVGVGLFLTPAR 3033
Db 3001 DIYHSVRARPRLLFGILLTFVGVGLFLTPAR 3033
RESULT 2
GNMVB8
genome polypeptide - hepatitis C virus (strain HC-J8)
N:Contains: capsid protein C; envelope protein M; hepatitis virus (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: A40250; F00397; F00559
R:Okamoto, H.; Kurei, K.; Okada, S.T.; Yamamoto, K.; Liuka, H.; Tanaka, T.; Fukuda, S.;
Virology 189, 331-341, 1992
A:Title: Full-length sequence of a hepatitis C virus genome having poor homology to repc
A:Reference number: A40250; MUID:92230232; PMID:1314459
A:Accession: A40250
A:Molecule type: genomic RNA
A:Residues: 1-3033 <OKA>
A:Cross-references: UNIPROT:P26661; GB:D10988; GB:D01221; NID:9221608; PIDN:BA01761.1;
R:Chan, S.W.; McMahon, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P.I
J. Gen. Virol. 73, 1131-1141, 1992
A:Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to e
A:Reference number: F00397; MUID:92268871; PMID:1316939
A:Accession: F00397
A:Molecule type: genomic RNA
A:Residues: 2678-2754 <CHA>
A:Cross-references: DDBJ:D10134
A:Experimental source: isolate E-B12
R:Kato, N.; Ootsubo, Y.; Ohnoh, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohno
Biochem. Biophys. Res. Commun. 181, 279-285, 1991
A:Title: Detection of plural HCV types in Japan.
A:Reference number: F00554; MUID:92068204; PMID:1720309
A:Accession: F00559

A:Molecule type: mRNA
A:Residues: 2678-2729 <KAT>
A:Cross-references: GB:D10567; GB:D090518; NID:9221523; PIDN:BA01418.1; PID:9221524
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructure
F:1-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <ME>
F:390-733/Product: nonstructural protein NS1 #status predicted <NS1>
F:734-1010/Product: nonstructural protein NS2 #status predicted <NS2>
F:1011-1619/Product: hepatitis virus #status predicted <NS3>
F:1234-1321/Region: nucleotide-binding motif A (P-loop)
F:1320-1323/Region: DEXH motif
F:1620-1866/Product: nonstructural protein NS4 #status predicted <NS4>
F:1867-2017/Product: nonstructural protein NS4b #status predicted <NS4b>
F:2018-3033/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,233,299,305,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,2038,23

Query Match 86.7% Score 13971; DB 1; Length 3033;
Best Local Similarity 83.9% Pred. No. 0;
Matches 2544; Conservative 245; Mismatches 244; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRBPQVKEPQGGQIVGVYLLPRGRGLGVRATRKTSERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRBPQVKEPQGGQIVGVYLLPRGRGLGVRATRKTSERSQPRG 60
QY 61 RROPTRKORRSRGKSGKRGYRPMPLVYNEGGLGMAWLLSPRGRSRSGWPNDRHRSRVNG 120
Db 61 RROPTRKORRSRGKSGKRGYRPMPLVYNEGGLGMAWLLSPRGRSRSGWPNDRHRSRVNG 120
QY 121 KYIDLTGSPADLMGYIPVGVAPLGVARALAHGVRVLEDGVNFATGNLPGCSFSIFLLA 180
Db 121 KYIDLTGSPADLMGYIPVGVAPLGVARALAHGVRVLEDGVNFATGNLPGCSFSIFLLA 180
QY 181 LLSCTTIVSAAEVXNISTGVWNTDCTNDSITWQLOAAVLHVPCCVPEKXGASQCI 240
Db 181 LLSCTTIVSAAEVXNISTGVWNTDCTNDSITWQLOAAVLHVPCCVPEKXGASQCI 240
QY 241 PVSBNVAVQRGALTOGRLTHIDVYMSATLCSALVYDGLCGVWLLAQMPLVSRQHMF 300
Db 241 PVSBNVAVQRGALTOGRLTHIDVYMSATLCSALVYDGLCGVWLLAQMPLVSRQHMF 300
QY 301 VODNCSTYPGTTHGRAMDMNMMSPTATMILYARVPEVIDIISGAMGVFGLA 360
Db 301 VODNCSTYPGTTHGRAMDMNMMSPTATMILYARVPEVIDIISGAMGVFGLA 360
QY 361 YFSMGANAKVAVVILLAAVDARTHTVGSAAQTGRLTSLFDMGPROKIQLVNTNSG 420
Db 361 YFSMGANAKVAVVILLAAVDARTHTVGSAAQTGRLTSLFDMGPROKIQLVNTNSG 420
QY 421 HINRPAALCNDLSHTGFIASLFYTHSFMSSGCEPMASACRSIARVVGALQYEDNTN 480
Db 421 HINRPAALCNDLSHTGFIASLFYTHSFMSSGCEPMASACRSIARVVGALQYEDNTN 480
QY 481 PEDMPPVGMHYPPRCGVASAKTVCGPYVCTPSPVVVGGTDRLAGAPTYMGENTDVL 540
Db 481 PEDMPPVGMHYPPRCGVASAKTVCGPYVCTPSPVVVGGTDRLAGAPTYMGENTDVL 540
QY 541 LNSTRPPLGSEFGCTMWNSSGYTKTCGAPPCRTADFNASTDLCTPCFCRGRPDYVLK 600
Db 541 LNSTRPPLGSEFGCTMWNSSGYTKTCGAPPCRTADFNASTDLCTPCFCRGRPDYVLK 600
QY 601 CGAGPMLTPRCLVDYDYLHMYPCVNVNPTIFARRYVGVEHRFSAACNFTRGDCRLED 660
Db 601 CGAGPMLTPRCLVDYDYLHMYPCVNVNPTIFARRYVGVEHRFSAACNFTRGDCRLED 660
QY 661 RDRSOLSPILSTTSMALIPCSYDLPALSTGLHLHONIYDVQVMYGSPLATYIYRM 720
Db 661 RDRSOLSPILSTTSMALIPCSYDLPALSTGLHLHONIYDVQVMYGSPLATYIYRM 720
QY 721 EWWILLFLLADARVCAQLMILLGQALAEKLVIIHAASASCSNGFLVFIFFVAAV 780
Db 721 EWWILLFLLADARVCAQLMILLGQALAEKLVIIHAASASCSNGFLVFIFFVAAV 780

Db 721 EWWILFLLLADARICACLMMLIILGQAEALKEKILILHSASASANGPIEFPTAAM 780
Qy 781 YIKGRVVPATYSLTGLMSEFSLLLALPOQAYAYDASVHGQIGALIMTTLTTPGX 840
Db 781 YLKRVRVPATYSLTGLMSEFSLLLALPOQAYALDAEGBELGALIVISITLTPAK 840
Qy 841 TLLSRFLMWLCYLLTLCBAMVQEMAPPMQVRGGRDGIIMAVAIFFGVFVDITKMLAVI 900
Db 841 ILSRSRWSMWSYMLVLAEOIQOVVPLFVRGGDGIIMAVAILHPLFVEVTKMLALIL 900
Qy 901 GPAYILKGLTRVYVFFRAHALILMCTMARHLAGRVVOMALLAGMGTGYIDHLTPM 960
Db 901 GPAYILKGLTRVYVFFRAHALILMCTMARHLAGRVVOMALLAGMGTGYIDHLTPM 960
Qy 961 SDMAAGLRDLAVABEPIIFSPMEKVIWGAETACGDIHGLPVASARLGRVLLGPD 1020
Db 961 STMAAGLRDLAVABEPIIFSPMEKVIWGAETACGDIHGLPVASARLGRVLLGPD 1020
Qy 1021 GYTSKMSLAPITAYAOQTRGLIGTIVSMTGKDTQAGEIOVLTSTQSTFLGTSIG 1080
Db 1021 GYTSKMSLAPITAYAOQTRGLIGTIVSMTGKDTQAGEIOVLTSTQSTFLGTSIG 1080
Qy 1081 VLMTVYHAGNKTLASGRPTOMVSSABGDIVGMPSPGTKSLBECTGADVLYITRN 1140
Db 1081 VLMTVYHAGNKTLASGRPTOMVSSABGDIVGMPSPGTKSLBECTGADVLYITRN 1140
Qy 1141 ADVIPARRGDKRGALSPRPLSTLKSGSGGPVLCPRGAVGVFRAVCSRGVAKSIDFT 1200
Db 1141 ADVIPARRGDKRGALSPRPLSTLKSGSGGPVLCPRGAVGVFRAVCSRGVAKSIDFT 1200
Qy 1201 PVETLIDVTRSPFSDNSTPPAVQTYOVGYLHAPTGSGSKTVPAVAAQYKVLVNP 1260
Db 1201 PVETLIDVTRSPFSDNSTPPAVQTYOVGYLHAPTGSGSKTVPAVAAQYKVLVNP 1260
Qy 1261 SVAATLFGAYLSKAGHINPIRTRVTVTTGAPITSTYKGFADOGCAGAYDIIICD 1320
Db 1261 SVAATLFGAYLSKAGHINPIRTRVTVTTGAPITSTYKGFADOGCAGAYDIIICD 1320
Qy 1321 ECHAVDSTTLIGTIVLDOAETAGVRLVLAATATPPGSVTTTPHNIIEVALGQGBEIPRY 1380
Db 1321 ECHAVDSTTLIGTIVLDOAETAGVRLVLAATATPPGSVTTTPHNIIEVALGQGBEIPRY 1380
Qy 1381 GRAIPLSYIKGGRLLTCHSKKKCDLAAALRGKGLNSVAYYGLDVSVPITGQDVVVA 1440
Db 1381 GRAIPLSYIKGGRLLTCHSKKKCDLAAALRGKGLNSVAYYGLDVSVPITGQDVVVA 1440
Qy 1441 TDALMTGYTGFDSVIDCNVAVOVVDFSLDPTFTITTOIOPDAYSRSORGRTRGRU 1500
Db 1441 TDALMTGYTGFDSVIDCNVAVOVVDFSLDPTFTITTOIOPDAYSRSORGRTRGRU 1500
Qy 1501 GIYRYVSTGERASGMPDSVYLCECYDAGAAWYELTPESETTVRLRAYFNTPEGLFVCDHLE 1560
Db 1501 GIYRYVSTGERASGMPDSVYLCECYDAGAAWYELTPESETTVRLRAYFNTPEGLFVCDHLE 1560
Qy 1561 FWEVAFVGLTHIDHFLSOTKQGENPAYITAYQATYCARAKAPPSMDVMKCLTRLKP 1620
Db 1561 FWEVAFVGLTHIDHFLSOTKQGENPAYITAYQATYCARAKAPPSMDVMKCLTRLKP 1620
Qy 1621 TLVGPFTLYRLSGVTNEVTLTHPVYKIATCMQADLEVTSTMPVLAGVLAAYVAYCA 1680
Db 1621 TLVGPFTLYRLSGVTNEVTLTHPVYKIATCMQADLEVTSTMPVLAGVLAAYVAYCA 1680
Qy 1681 TGCVCIIIGRLHINORAVAPDKEVLYBAFDEMECASAALIEGQRIABMLSKIQGLL 1740
Db 1681 TGCVCIIIGRLHINORAVAPDKEVLYBAFDEMECASAALIEGQRIABMLSKIQGLL 1740
Qy 1741 QOASKOADOTOPYQOASMPKVEOFMAGMNFISGIOYLAGLSTLRBNPVAASMAFSA 1800
Db 1741 QOASKOADOTOPYQOASMPKVEOFMAGMNFISGIOYLAGLSTLRBNPVAASMAFSA 1800
Qy 1801 LTPSLSTSTIILNIGMGLASQIAPAPAGATGFVSGVGAAGVSGIGLGVLDIILAGV 1860
Db 1801 LTPSLSTSTIILNIGMGLASQIAPAPAGATGFVSGVGAAGVSGIGLGVLDIILAGV 1860

Qy 1861 AGISGALVAFKIMSGEKSPMEDVVNLIPGLISPGALVVGVICAAILRRHVPGEGAVQM 1920
Db 1861 AGISGALVAFKIMSGEKSPMEDVVNLIPGLISPGALVVGVICAAILRRHVPGEGAVQM 1920
Qy 1921 NRIIFASRGHNVATHTYVTSASQRYTOLLGSLITLSLRLLHMTTEBCPIPCGSM 1980
Db 1921 NRIIFASRGHNVATHTYVTSASQRYTOLLGSLITLSLRLLHMTTEBCPIPCGSM 1980
Qy 1981 LRDVMDWCTILITDKKMLTSLPROMGLFPVSGQKYGKVMAGTGMTRCPGANAIS 2040
Db 1981 LRDVMDWCTILITDKKMLTSLPROMGLFPVSGQKYGKVMAGTGMTRCPGANAIS 2040
Qy 2041 GNVRLGSRNITGPKTCMNIWQTFPINCYTEGQCQCPKAPNFKVALMFAASEVAYTOH 2100
Db 2041 GNVRLGSRNITGPKTCMNIWQTFPINCYTEGQCQCPKAPNFKVALMFAASEVAYTOH 2100
Qy 2101 GSYYITGLITDNLKVPQOOLSPSEFFSVWDGVOIHRPAPTRKPPFRDVSFCVGLNSFV 2160
Db 2101 GSYYITGLITDNLKVPQOOLSPSEFFSVWDGVOIHRPAPTRKPPFRDVSFCVGLNSFV 2160
Qy 2161 GSQLPDPEPPTDVLMSMLTDPSSHITARTAAARLARGSPSASASASQLSAPSLARTCT 2220
Db 2161 GSQLPDPEPPTDVLMSMLTDPSSHITARTAAARLARGSPSASASASQLSAPSLARTCT 2220
Qy 2221 THKAYVDVMDVANDLFWMGDVTTRIESGSKVYVULSLDPMVERSDLPSPISBYMLPKR 2280
Db 2221 THKAYVDVMDVANDLFWMGDVTTRIESGSKVYVULSLDPMVERSDLPSPISBYMLPKR 2280
Qy 2281 FPPALPAMARPDYNPPLVESKRPDYOPATYAGCLAPPRAKTPPPPPRRARTVGLSEST 2340
Db 2281 FPPALPAMARPDYNPPLVESKRPDYOPATYAGCLAPPRAKTPPPPPRRARTVGLSEST 2340
Qy 2341 GVALQOLAIKSRGQPPSGDSGLSTGAGAAUGSGSTPDEBLALSTGSISSMPLEBGL 2400
Db 2341 GVALQOLAIKSRGQPPSGDSGLSTGAGAAUGSGSTPDEBLALSTGSISSMPLEBGL 2400
Qy 2401 DEDLEPBOVEPOPPEQGVAAFGSDSGSWTSCSEBDSVWCSSMSYWTGALITPCSEBE 2460
Db 2401 DEDLEPBOVEPOPPEQGVAAFGSDSGSWTSCSEBDSVWCSSMSYWTGALITPCSEBE 2460
Qy 2461 EKLPTNPISNLSLRHNVYCTTYSASLRKAKTTFDMOVLDSYDVSVDKILAAKV 2520
Db 2461 EKLPTNPISNLSLRHNVYCTTYSASLRKAKTTFDMOVLDSYDVSVDKILAAKV 2520
Qy 2521 TARTLTMEBACOLTPPHSARSKYGFAGKEVNSLSGRAVNHILKSWKDLJDSSEPIPTI 2580
Db 2521 TARTLTMEBACOLTPPHSARSKYGFAGKEVNSLSGRAVNHILKSWKDLJDSSEPIPTI 2580
Qy 2581 MAKNEVFCVDPPTKGKKAARLIYVPLDGVRCERKALYDITOKLPOAVMGASYGQYSPA 2640
Db 2581 MAKNEVFCVDPPTKGKKAARLIYVPLDGVRCERKALYDITOKLPOAVMGASYGQYSPA 2640
Qy 2641 QVREFLLKMAEKDQPMGFSDTRCFDSTYVERDRTREESLYOAGSLQOEARLYHSLTE 2700
Db 2641 QVREFLLKMAEKDQPMGFSDTRCFDSTYVERDRTREESLYOAGSLQOEARLYHSLTE 2700
Qy 2701 RLYVGGPMFNSKQTCGRRCRASAGVLTSSMGNITTCVKAALACKAAGIAPTMLVCGD 2760
Db 2701 RLYVGGPMFNSKQTCGRRCRASAGVLTSSMGNITTCVKAALACKAAGIAPTMLVCGD 2760
Qy 2761 DLVVISSEQTEEDERNIRAFTEAMTRYAPPGDPREPYDLELITSCSSNVSAVALGPOG 2820
Db 2761 DLVVISSEQTEEDERNIRAFTEAMTRYAPPGDPREPYDLELITSCSSNVSAVALGPOG 2820
Qy 2821 RRRYVLTDPPTPIARAAMTVRHSPVNSWIGNIIQVAPTIWVRVINTHFFSILLADT 2880
Db 2821 RRRYVLTDPPTPIARAAMTVRHSPVNSWIGNIIQVAPTIWVRVINTHFFSILLADT 2880
Qy 2881 LDONLNFEMGAVVSVPLDPAITERLHGLDAFSLHTYTTHBELTRVSALEKLGAPLR 2940
Db 2881 LDONLNFEMGAVVSVPLDPAITERLHGLDAFSLHTYTTHBELTRVSALEKLGAPLR 2940

QY	1381	RRAIPPLSYIKGGRNHLIFCHSKKKCODELAALRMGLNSVAAYYRGIDSVIPFGODVYVA	1440
Db	1377	GKAIPLEIAIKGGRNHLIFCHSKKKCODELAAKLTGELNAAVYRGIDSVIPFGSDVYVA	1436
QY	1441	TDALMTGYGPFDSVIDCNAVAVTQVDPDSLDPFTLITQIYIPODAVSRQGRGTGRL	1500
Db	1437	TDALMTGTGPFDSVIDCNCVTQVDPDSLDPFTLITLTPRLQAVSRQGRGTGGRS	1496
QY	1501	GIYRVYSTGERASGMFDSVILCEYDAGAAVYELTPESTTVRLAVENTPGLPYCODHLE	1560
Db	1497	GIYRVFTTGERBPSGMFDSVILCECYDAGAAVELTPAETSVRRLAVENTPGLPYCODHLE	1556
QY	1561	FMEAVFTGLTHDAHFLESOTKOSGNGFAYLTAQVATVCAKAPRPSDMMWKCLTFLKP	1620
Db	1557	FMESVFTGLTHDAHFLESOTKQAGDNLPYLAQVATVCAKQAPPSPDQMKCLIRLKP	1616
QY	1621	TLVGPFTPLLYRLSGVSTNEVTLTHPVTKYIAVCMQADLEVMSTWLAGVLAATAAACL	1680
Db	1617	TLHGFTPLLYRLSGAVQNEVTLTHPTTKYIMCMSADLEVTSTWLVQGVLAATAACLT	1676
QY	1661	TGCVCITIGRLHINQRAVVAIPDKEVLYEAFDEMEBCASPAALIEGQRIAEMLKSKIQGLL	1740
Db	1677	TGSVYIVYGRITILSGRPVAPIDREVLYQGFDEMEBCASHLPYIEQMQIAQEPFKQKALGLL	1736
QY	1741	QOASQVQADIDQPTVOASMPKYBQFPAAKMMNFISGIQYLAQSLTLPGRPAVASMAFSA	1800
Db	1737	QVATQVQAAAPAVVESKRALEVEFPAAKMMNFISGIQYLAQSLTLPGRPALIASLMAFTAS	1796
QY	1801	LTPSLPSTTLLNLNIGMLASQIAPAPGAGVPGVSGVGAANSIGIGKLVYLLAGYG	1860
Db	1797	ITSPLTTONTLIFNLNIGMWVAAQIAPAPASAPFAGAGIAGAANSIGIGKLVYLLAGYG	1856
QY	1861	AGISGALVAFKIMSGEKPSMEDVYNLIPGLISPGALVGVICAAILRRHNGPBGGAQVM	1920
Db	1857	AGVAGALVAFKIMSGEMSTEDVYNLIPALISPGALVGVICAAILRRHNGPBGGAQVM	1916
QY	1921	NRLIAFASRGNHVAPTHVYTESDSQRTQOLGSLTTSLLRLHNMITDECPIPCGGSW	1980
Db	1917	NRLIAFASRGNHVSPTHVYTESDAAARTQOLISLITQILKRIHOWINEDCSPCGGSW	1976
QY	1981	LRDVDMVCTILTDPKMLTSLKLPKMGFLPFVSQOKYKGVMGGTGIMTRPCGANSIS	2040
Db	1977	LKDVMDMICTVLSDFKTLQSKLPRLEGLPEPLSCQKQYKGVMGDGMQTTCCGQAQIT	2036
QY	2041	GNVRIGSNRITGPKTCMMNIWQSTPEINCYCTEQOCQPKAPMPFKAIWRVASEVAYTOH	2100
Db	2037	GHVKGSGMRIYGPKTCSTNMHTGTPINAYTTGQCPSPAPARYSALMRVAAEEVETRV	2096
QY	2101	GSYHYITGLTDNLKVPQCLPSPPEFFSWDQVQHRPAPEPKPPRDESVSCVGLNSFV	2160
Db	2097	GDPHYVTGMITDNYKCPQVAPAEFEFTVDGVRILHRYPVCKPLIREVVFQVGLNQYLV	2156
QY	2161	GSOLPCDEBPDPVLMSTLDPSSHITATARRLARGSPEBASASSASQLSAPBLATCT	2220
Db	2157	GSOLPCDEBPDPVAVLTSMSTLDPSSHTATARRRLARGSPELASSASQLSAPBLKATCT	2216
QY	2221	THGKAYVDVDMYANLE-----MGADYTRIESGSKVVLVDSLDPMWEEBSDEPSI:PSEXML	2276
Db	2217	THHSRPAIDLLEANLIMQEMGNGITTRVESENKVIILDSPPRIAVDEBERISYPAELIR	2276
QY	2277	PKKRPFPALPMWARDYNNPPLVESMKRDPQAPATVAGCALPBRPKTPTPPRRRRYGLS	2336
Db	2337	ESTVSSALAEALATKTFG-----SSGSASAVDSGTATPBPQASDDGDKGDVDSYS	2386
QY	2391	SMPPLEGELGPDLEPEQVEQPPQGGVAAFGSDSGMSTCS--BEDDSVCCGMSYSWT	2449
Db	2387	SMPPLEGEPGPDL-----SDGMSVYSGBAGDVVCCGMSYWT	2426
QY	2450	GALITPCSPBEBKLPINELSNSLLRHNKVVCTTTKASLRAKVTFPRMQVLDYYDSV	2509

Query Match 74.3%; Score 11984.5; DB 1; Length 3010;

Best Local Similarity 71.5%; Pred. No. 0;
Matches 2175; Conservative 352; Mismatches 472; Indels 45; Gaps 8;

QY 1 MNTNRPQAKTKRNTNRPRQDVYFPGGGQIVGGVYLLPRGRPLGRATRKTSERSQPRG 60
Db 1 MNTNRPQAKTKRNTNRPRQDVYFPGGGQIVGGVYLLPRGRPLGRATRKTSERSQPRG 60
QY 61 RQOPIDKDRSTGSKGPKGYPMPLYGNBGLMAGMLSPGSRPSMGNDPRHRNRNG 120
Db 61 RQOPIDKARRRPRGRRTAOPGYPMPLYGNBGLMAGMLSPGSRPSMGNDPRHRNRNG 120
QY 121 KYIDLTGCFADLMGYIPVVGAPLGAGVABALAHGVLEDDGVNFATGMLPGCSFSIFLLA 180
Db 121 KYIDLTGCFADLMGYIPVVGAPLGAGVABALAHGVLEDDGVNFATGMLPGCSFSIFLLA 180
QY 181 LLSCTTPVSAABVKXISTGVMTNDCTDSTWQLOAAVLHVPGCVPEKQGNMSQCM 240
Db 181 LLSCTTPVSAABVKXISTGVMTNDCTDSTWQLOAAVLHVPGCVPEKQGNMSQCM 240
QY 241 PVSNAVAPRGALPGALRTHIDMVYMSATLCSALVVGDLGGVMAAQMFIYSPQHMF 300
Db 241 ALTPILARNVTIPTTIRRHVDLVGAALFCSAMVGDLCSSVFLVSQULFTSPRRHYT 300
QY 301 VODNCNSIYPTGITGHRMAMDMMMNSPTATMILAYAMEVPEYIDIIISGAMGVNFGLA 360
Db 301 LQDNCNSIYPGHVSGRMAMDMMMNSPTATLVVSQLARIPOAVDVMAVGAHGVYLAGLA 360
QY 361 YFSMAGAKKVVILLILAAGVDARTHTVGGSAQTTGRLTSLFDNGPROKIQLVNTNGW 420
Db 361 YFSMAGAKKVVILLILAAGVDARTHTVGGSAQTTGRLTSLFDNGPROKIQLVNTNGW 420
QY 421 HINRTALNNDLSHGEFIALFYTHSFNSGGERBMSACRSIEMPVGALQOYEDNVN 480
Db 421 HINRTALNNDLSHGEFIALFYTHSFNSGGERBMSACRSIEMPVGALQOYEDNVN 480
QY 481 PEDMBPYCMHYPPOGCVVSAKTVCGPVYCTPSPVVGTTRDLGAPYTWGSENETDVL 540
Db 479 RSDQRPYCMHYPPOCTIVPASEVCGPVYCTPSPVVGTTRDFGVPTRKMEBNETDVL 538
QY 541 LNSTRPGLSMPGCTMWNSSGYTKTCGAPPCRTADFNASTDLCEPTDCFRKHPDITYLK 600
Db 539 LNSTRPGLSMPGCTMWNSSGYTKTCGAPPCRTADFNASTDLCEPTDCFRKHPDITYLK 596
QY 601 CGSGPMLPRCLIDVPRYLMHYPCMYNTVIFKIMVGVGEVHRILTAACNFTGDRONED 660
Db 597 CGSGPMLPRCWDVYPRYLMHYPCMYNTVIFKIMVGVGEVHRILTAACNFTGDRONED 656
QY 661 RDRSOLPLSHSTTEWAILPCSYDLPALSTGLIHLHONIVDQFMYGLSPALTKYIVM 720
Db 657 RDRSOLPLSHSTTEWAILPCSYDLPALSTGLIHLHONIVDQFMYGLSPALTKYIVM 716
QY 721 EKVYLLFLLADARVACACIMMLILQAQEALEKLVILHAASACNGFYVIFVVA 780
Db 717 EKVYLLFLLADARVACACIMMLILQAQEALEKLVILHAASACNGFYVIFVVA 776
QY 781 YIKGVVPLATYSLINGMFSILLALPOAAVAYDASVHGOIGALLVNTLEPTLRGK 840
Db 777 YIKGVVPLATYSLINGMFSILLALPOAAVAYDASVHGOIGALLVNTLEPTLRGK 836
QY 841 TLLSRFLMWLCYLLTLGEAMVOEMAPPMQVGRGDDIWAVALFYEGVVFDTTKMLAVL 900
Db 837 VELALUIMWLYFTTRAEADLHVMIPLNARGRDAILLMCAVHPELIDITKLIALL 896
QY 901 GPAYILKGLATVPYFVRAHALRMCCTMARHLAGRYVOMALLAIGRMVTGYIYDHLTM 960
Db 897 GELMYLQAGITRVPYFVRAQGLIHACMLVRKVAAGGYVQMAFMKIGALGTIYVHMLTFL 956
QY 961 SPMAASGLBDLVAVEPIIFSPMEKKVIYGAETAAGDIIHGLPVSAALGSEVLGPD 1020
Db 957 RQMPRAGLBDLVAVEPIIFSPMEKKVIYGAETAAGDIIHGLPVSAALGSEVLGPD 1016
QY 1021 GYTSKMSLLABITAYAQOTRGLGTIVVSMTRDTEQAGEIOVLSTVOSFLGTSISG 1080
Db 1021 GYTSKMSLLABITAYAQOTRGLGTIVVSMTRDTEQAGEIOVLSTVOSFLGTSISG 1080

Db 1017 SLEGRGLRLAPITAYSOQTRGLCIITSLTGRDNQVEGEVQVSTATQSFATCVNG 1076
QY 1081 VLVMTYVHAGNKTLAGSGRPVYQWYSSABGDLVGMSPSPGTSLEPCCGANDLVLYRN 1140
Db 1077 VCMTVYHAGSKTTLAAPGPIYQWYVNDQDLVGMKPKPGASLTIPCCTGSSDLVLYRH 1136
QY 1141 ADVIPARRRGDRGALLSPRLSTLKGSSGGFVLCPRGHVAVGFPAVAAVSRVAAISDPI 1200
Db 1137 ADVIPARRRGDRGALLSPRLSTLKGSSGGFVLCPRGHVAVGFPAVAAVSRVAAISDPI 1196
QY 1201 PVEITLIDVTRSPFSDNSTPPAVPQTYGVYLAHPTGSGSKTVKPVAAVAGYKVLVLP 1260
Db 1197 PVESEMTMRSPVFTDNSSPPAVPOSFOVAHLHAHPTGSGSKTVKPVAAVAGYKVLVLP 1256
QY 1261 SYAATLGGAYLSKAGHNPINRTGVRVTGGAPIYTSYTGKFLADGGCAGAGAYDIIID 1320
Db 1257 SYAATLGGAYLSKAGHNPINRTGVRVTGGAPIYTSYTGKFLADGGCAGAGAYDIIID 1316
QY 1321 ECHAVDSFTIIGITVLQDAETAGVRLTVLATATPPGSGVTTPHPNIEVALQOGEIIPY 1380
Db 1317 ECHAVDSFTIIGITVLQDAETAGVRLTVLATATPPGSGVTTPHPNIEVALQOGEIIPY 1376
QY 1381 GRAIPLSYIKGGRHLIFCHSKKKCCDELAALRGKLSNVAAYYRGLDVSIPTQGDVVVA 1440
Db 1377 GRAIPLEAIKGRGRHLIFCHSKKKCCDELAALRGKLSNVAAYYRGLDVSIPTQGDVVVA 1436
QY 1441 TDALMTGCTGDPDSYIDCNVAVTOYVVDPSLDPTFTTQIYQVQDASVRSORGRGRGL 1500
Db 1437 TDALMTGCTGDPDSYIDCNVAVTOYVVDPSLDPTFTTQIYQVQDASVRSORGRGRGL 1496
QY 1501 GYRYVSTGERASGFPDVSIVCECYDAGAAWELTPESETTVLRAVFTPGLPVQDHL 1560
Db 1497 GYRYVSTGERASGFPDVSIVCECYDAGAAWELTPESETTVLRAVFTPGLPVQDHL 1556
QY 1561 FWEAVFTGLTHIDAHFISQTKQSGENFAYLTAQATVCARAKAPPSNDVMMKLTRLKP 1620
Db 1557 FWEAVFTGLTHIDAHFISQTKQSGENFAYLTAQATVCARAKAPPSNDVMMKLTRLKP 1616
QY 1621 TLVGPPLLYRLGASTNEVTLTHPYTKYIATCMQDLSEVMTSTWVLAGVLAANAAYCLA 1680
Db 1617 TLVGPPLLYRLGASTNEVTLTHPYTKYIATCMQDLSEVMTSTWVLAGVLAANAAYCLA 1676
QY 1681 TGCVCIGRLHINOAAVAPDEKLYEAPDEMEECASRAALIEBQRIAEMLKSIQGL 1740
Db 1677 TGCVCIGRLHINOAAVAPDEKLYEAPDEMEECASRAALIEBQRIAEMLKSIQGL 1736
QY 1741 QOASHQADIDPTVQASWPKVEQFQWAKMMNFISGIOYLAGLSTLPGNPVAVMAAFSA 1800
Db 1737 QOASHQADIDPTVQASWPKVEQFQWAKMMNFISGIOYLAGLSTLPGNPVAVMAAFSA 1796
QY 1801 LTPSLSTTILINLIGWMLASQIAPPAGATGVVSGLVGAAGVSGIGLKVLDVTLAAG 1860
Db 1797 LTPSLSTTILINLIGWMLASQIAPPAGATGVVSGLVGAAGVSGIGLKVLDVTLAAG 1856
QY 1861 AGISGALVAFKIMSGEKSPMEDVNLPGILSPGALVVGVCALILRRHVGESEAVQWM 1920
Db 1857 AGISGALVAFKIMSGEKSPMEDVNLPGILSPGALVVGVCALILRRHVGESEAVQWM 1916
QY 1921 NRLIAFASGRNVAPTHVYTESDAQRYVTOGLGSLTISLRLRLNMVITTEDCPICGSG 1980
Db 1917 NRLIAFASGRNVAPTHVYTESDAQRYVTOGLGSLTISLRLRLNMVITTEDCPICGSG 1976
QY 1981 LRDVDMVCTIITDPRKMLTSLPFKMGCLPVSQKXKGYMAGTGIMTRCPGAGNIS 2040
Db 1977 LRDVDMVCTIITDPRKMLTSLPFKMGCLPVSQKXKGYMAGTGIMTRCPGAGNIS 2036
QY 2041 GNVRLSGKRITGPKICMNIWQSTPINCYTGQOCPKPAKPVKVAIMVAASEVLEVTQH 2100
Db 2037 GNVRLSGKRITGPKICMNIWQSTPINCYTGQOCPKPAKPVKVAIMVAASEVLEVTQH 2096
QY 2101 GSYHITGLTLDNLKVPQCLPSPEFFSVVDVQVHRAFPKPFPRDVSFCVGLNSFV 2160
Db 2097 GSYHITGLTLDNLKVPQCLPSPEFFSVVDVQVHRAFPKPFPRDVSFCVGLNSFV 2156

QY 2161 SSOLPCDEPPTDVLSMLTDPSSHITTAATAARRLARSGSPSEASSSAQLSPASIRATCT 2220
 Db 2157 GSOLPCEBEPDVAVLTSMLTDPSSHTAETARRRLARSGSPSIASSSAQLSPASIRATCT 2216
 QY 2221 THGKAYDVMDVANDANF-----MGADVTRIESGSKVVLVLSLPMVERSDLESPISSEYML 2276
 Db 2217 THHSPPADLLIENLILMRQEMOGNITRVESENKVVVLSFDPRLAREBEREVSVAELLR 2276
 QY 2277 PKGRFPALPAMARPDPNPLVESWKRPDYQAVVAGCALPPPRKTPPPRRRTVGLS 2336
 Db 2277 KSKKFPAMPIWARPEDYNPPLIESWKDDYVPVHGCPLPIKAPPIPPRRRKTVVLT 2336
 QY 2337 EDSIGDALQOLAIKSFQGPSPSGSGSLSTGAADSPSGOTP--PBLA-----LSETGIS 2330
 Db 2337 ESSVSALAEALTKTFG-----SSBASVDSGATATLPDASDDGKGSVEYS 2386
 QY 2391 SMPLELEGDPDLEPEQVEPQPPQGGVAAFGSDSGMSTCSER--DDSVVCCSMYSWT 2449
 Db 2387 SMPLEEGRPDDL-----SDGSMVSYSEBASDVCCSMYSWT 2426
 QY 2450 GALTTPCSPEBEKLPINPLNSLRYHNKVCITTKASLRKATYEDRMQVLDYSYDSV 2509
 Db 2427 GALITPCAAEESKLPINALNSLIRHNMVVATTSRSGLRKQKVPRLQVLDHYRDV 2486
 QY 2510 LKDIILASXYTARLLTMEBAQCLTPPHSASKSYFGAKVRSISGRVNHIIKSWKDL 2569
 Db 2487 LKEMAKASYAKVAKLLSVEACKLPHPHSASKFYGAKDVRRNLSKAVNHHSWKDL 2546
 QY 2570 EDSEPTPIITTMAKNEVCVDPPTKGKKAARLIIVPDIAGVNCESKALYDITOKLPQAVM 2629
 Db 2547 EDTVPIPIDTITMAKNEVFCVQPEKGRKPARLIIVPDIAGVNCESKALYDITOKLPQAVM 2606
 QY 2630 GASVTFQYSPARVFEFLKMAEKDPMGFYDITCFDSTYTERDIRPEESIYRACSLPE 2669
 Db 2607 GSSVGFQYSPGQRFVEFLVNTWKSQKPMGFSDYTCFSDSTYENDIRVEESIYQCCDLAP 2666
 QY 2690 EAHTIHLTEELVYVGGMENSKGOTCGYRCRASAGLTSMGNITTCYVVALAACRAG 2749
 Db 2667 BRQAKIKLTELVLIGPLVMSKQNGCGRRRAGVLTSCGNLTTCYVALAASACRAK 2726
 QY 2750 IIAPTMLVCGDVLVVISSEQTEEDERNLRAFTAMTRYSAAPGDPPEPEYDELITSCS 2809
 Db 2727 LQDCMLVNGDVLVVICESAGTQEDASLRVTEAMTRYSAAPGDPPEPEYDELITSCS 2786
 QY 2810 SNVSVALPQGRRYLITRDPTPIARAAMETVRSVNSWLGNTIYAAPITMAMVMT 2869
 Db 2787 SNVSVAHDASGRVYLLTRDPTPIARAAMETARTPNVSWLGNTIYAAPITMAMVMT 2846
 QY 2870 HFFSILMADDTLDOLNFMFYGAIVSVSLDPAIIRLHGLGDAESLTYPHLETRVVS 2929
 Db 2847 HFFSILLMOBOLKRLDCIYACYSIEPLDPOIIRHGLSABLSHSISGEENRVAS 2906
 QY 2930 ALRLKGAAPPLRAMKSARAVRASLISRGRAAVCGRYLFNNAVKTKLKLTPLEARLIDL 2989
 Db 2907 CLRKLGVPPLRVWRHARSVRRARLLSCGGRATCGKYLFNNAVKTKLKLTPLEASRLDL 2966
 QY 2990 SSMFTVAGGGDIYHSVRARPRLLIFGLILLFVAGLFLIPAR 3033
 Db 2967 SGMFVAGYSGDIIYHLSRRARPRMFLCLILLVAGVGIYLLPNR 3010

RESULT 5
 A45573
 genome polyprotein - hepatitis C virus (strain J7)
 N:Contains: capsid protein C; envelope protein M; hepatitisin (BC 3.4.21.98) (nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
 C:Species: hepatitis C virus
 C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
 C:Accession: A45573
 R:Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.-J.; Nakazawa, T.; Hijikata, Y.; Iwano, R.; 23, 39-53, 1992
 Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: s

A:Reference number: A45573; MUID:92295714; PMID:1186627
A:Accession: A45573
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3010 <RAN>
A:Cross-references: UNIPROT:Q00269; GB:D11168; GB:D01171; NID:g221612; PIDD:BA01943.1;
A:Experimental source: HCV-JT
A:Note: sequence extracted from NCBI backbone (NCBN:106206, NCBIP:106207)
C:Superfamily: hepatitis C virus genome polypeptin
C:Keywords: ATp; glycoprotein; hydrolase; nucleotide binding; P-loop; polypeptin; serin
F:2-115/Product: capsid protein C #status predicted <PC>
F:116-191/Product: envelope protein M #status predicted
F:199-369/Product: major envelope protein E #status predicted <ME>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: nonstructural protein NS3 #status predicted <NS3>
F:1220-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4 #status predicted <NS4>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <NS4b>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>

QY	I	MSTNPKRQKTKRNTNRPRQDYKFPFGGQIYGVGYLLPRGRPLGVPAKTKTSERSOPRG	60
Db	1	MSTNPKRQKTKRNTNRPRQDYKFPFGGQIYGVGYVLLPRGRPTLGVPAKTKTSERSOPRG	60
QY	61	RRQPIFKDRSRGKSWKGRPYMPRLYXNEGIGMAQWLLSPGSRPSMGPNDPHRSRNVG	120
Db	61	RRQPIFKARPRGRAMAQPGYMPRLYXNEGIGMAQWLLSPGSRPSMGPTDPRRRSNNG	120
QY	121	KVIDTLTCGFADLMGTYIPVVGAPLGVAPALAHGVRVLEEDVNATGNLPGCSFSIFLLA	180
Db	121	KVIDTLTCGFADLMGTYIPVVGAPLGVAPALAHGVRVLEEDVNATGNLPGCSFSIFLLA	180
QY	181	LISCITTPVSAAEVKNIISTGYVTVNDCTNDSITWQLOAAVLHRCVPCSEKVGNAQOCWT	240
Db	181	LISCITTPASAYOVRAAGSLYHVTNDCNSSIYVEAAGMIHNTGCVPCRENNASCMW	240
QY	241	PVSPNVAOVRGALNQGRLTHIDVVMVATCSALYUGDLCGGWMLAOMFVTSPOHNHF	300
Db	241	ALPPLTAARTSIPITTIKRHVDLIVGAAPCSAAYUGDLGGSVFLVYSQLTSSPRYET	300
QY	301	VQDQNCISYPGTITIGHRAMDMNMNMNSPTAIIILAYARVEVLIIDISGAHGVMEGLA	360
Db	301	VQDQNCISYPGHVSGRHAMDMNMNMSPTLALVYSQLLRIPQAVDVMAHMGVLAGLA	360
QY	361	YFSQMGAKAVVILLILAAQVDAPIHYVYGGSAAOQTGLTSLFPMGRQKIQLVNTYNSW	420
Db	361	YFSVQVMAKAVILMLLFAQVDGVYTYTGGQARITQSVTSFFPQGAORIQOLINTNSW	420
QY	421	HIKRTALNCDSLHTGFIASLFPYTSFPMSSGCPERMASCRSIEAFRGMGALOYEDVNM	480
Db	421	HIKRTALNCNESLNTGFPALFYAAKFNSSGCPERMASCSIDKFAQMGKPIYY---TE	476
QY	481	PEDM--RPYCMHYPRROCGVVSAKTVCGPVYCFPPSPVVGTTDRGLAPTYWGENETDV	538
Db	477	PRLDQRPYCMHYAPROGCIIVPAQVQCPVYCFPPSPVVGTTDRSAPLYNMGANETDV	536
QY	539	FLINSTRPPLGSGMFCGTMMNSSGYTKGCAAPRCRTADFNASTDLLPPTDCFRKHPTTY	598
Db	537	LILNTRPRPGQWNVFGCTMMNNSGTGFKTCGCPPCNIGGVNGL--LTCPTDCFRKHPEATY	594
QY	599	LKQSGGPMILPRCLIDYPRYLHMYCTVNYTIFKIRMYGVGVERRLLAAACNFRGDCNL	658
Db	595	TKQSGGPMILPRCLIDYPRYLHMYCTVNYTIFKIRMYGVGVERRLLAAACNFRGECDL	654
QY	659	EDKDRQSLPLLHSTTEWAILPCSSYDLPALSTGLLHNOIIVQVQMGSLPALTYKIV	718

Db 655 EDRDRSELSPLLLSTTEMQTLPCSEFTTLPALSTGLIHLHQNIVDQYLYIGISAVVSFVI 714
 Qy 719 RHEWVILFLIADARVACCLMMLLLQGAELKVIYLHAAASAACNGFYPIYFPA 778
 Db 715 KMEYVLLFLLIADARVACCLMMMLLIAQEALENLVNLHAAASLAGAGLISFVFPFA 774
 Qy 779 AMYIKGRVPLATYSYLTGLMSFSLLLALPOQAYAVDASVHGOIGALLVMTLFTLTPG 838
 Db 775 AMYIKGRVPLATYSYLTGLMSFSLLLALPPRAYAMDREMAASCGVVFVGLILTLSPH 834
 Qy 839 YKTLISRLPLMWLCYLLTLGEANVOEMAPMOVGRGDGIYMAVAIPYGVFDTIKMLIA 898
 Db 835 YVFLARLILMWLYQYITBAEHLQVWVPLNRCGRDAIILTLCAHPELIDITKLLIA 894
 Qy 899 VLGPAVYLKGALTRPYFYRAHALLMCMTHARHLAGRYVQWALLALGMCWTYIYDHT 958
 Db 895 ILGPMVLIOALITAMPYFVRAQGLIRACLVKRVAGHYVQMAFKMLAALTETYYDHT 954
 Qy 959 PMSDMAASGLRDLAAVEPIIPSMBKVIYVMAETACGDILHGLPVASARLGRVLLGP 1018
 Db 955 PLODMAHAGRLDAAVEPVVSDMETKITWGLADTACGDILGLPVASARLGRVLLGP 1014
 Qy 1019 ADGYTSKMSLLAPITAYAOQTRGLLGTIVSMTGRDKTEAGEIQVLSVTYQSEFLGSI 1078
 Db 1015 ADSIGOGWRLLAPITAYAOQTRGLLGCIVTSLTGRDKNOVGEVQVSTATQSFILATCV 1074
 Qy 1079 SGVLTMYHAGNKTLAGSRGVYQWYSSAEGDLVGMSPSPGTSLEPCTGCAVULYVT 1138
 Db 1075 NGCVMTVHFAGSKTLAGKGPITQMYTNVDOLVGMHAPGARBSLTPTCTCSSDLXYLT 1134
 Qy 1139 RNADVIYPARBRDCKGALLSPPLSTLKGSSGSPVLCPRGHAVGVFRAVAVCSRGVAKSID 1198
 Db 1135 RHADVIYPARBRDCKGALLSPPVSYLKGSSGSPVLCPRGHAVGVFRAVAVCSRGVAKSID 1194
 Qy 1199 FIPVETLIDIVRSPTFSDNSTPBAVPOYQVGYLAHPTSGSKSTYVPAVAAQYKVLVLT 1258
 Db 1195 FIPVSMETVHRSPTFSDNSTPBAVPOYQVHNLHAPTSKGSKTKVPAVAAQYKVLVLT 1254
 Qy 1259 NPSVAATIGFGAYLSKAGININITGVRTVTTGAPITYSTYKGLADGCGAGAYDITI 1318
 Db 1255 NPSVAATIGFGAYLSKAGININITGVRTVTTGAPITYSTYKGLADGCGAGAYDITI 1314
 Qy 1319 CDECAVNSTTILIGITVLDQAEAGVRLTVLATAATPPGCVTTTPBNIEEVALGQGEIIP 1378
 Db 1315 CDECAVNSTTILIGITVLDQAEAGVRLTVLATAATPPGCVTTTPBNIEEVALGQGEIIP 1374
 Qy 1379 FYGRAIPLSYIKGGRHLIFCHSKKKCCDELAALRGWGLNSVAYYRGDLVSVIPTQDVV 1438
 Db 1375 FYGRAIPLSYIKGGRHLIFCHSKKKCCDELAALRGWGLNSVAYYRGDLVSVIPTQDVV 1434
 Qy 1439 VATDALMTGYTGDPPSVIDCNVAVTQVYVDFSLDPTFTTTTQIYVPODASRSQRGRG 1498
 Db 1435 VATDALMTGYTGDPPSVIDCNVAVTQVYVDFSLDPTFTTTTQIYVPODASRSQRGRG 1494
 Qy 1499 RLGIRAYVSTGBRSGMFDPSVYLGCYDAGAAWYELTSETTVRLRAYFNTFGELVCDH 1558
 Db 1495 RLGIRAYVSTGBRSGMFDPSVYLGCYDAGAAWYELTSETTVRLRAYFNTFGELVCDH 1554
 Qy 1559 LEFMEAVTGLTHIDAHFLSQTQSGENFAYLYTAQATVCARAKAPPSMDVMKCLTRL 1618
 Db 1555 LEFMEAVTGLTHIDAHFLSQTQSGENFAYLYTAQATVCARAKAPPSMDVMKCLTRL 1614
 Qy 1619 KPTLVGPPLLYRLGSLVNEVTLTHPVTKYIATCMQADILEVMTSTWVLAAGVLAANAAC 1678
 Db 1615 KPTLVGPPLLYRLGSLVNEVTLTHPVTKYIATCMQADILEVMTSTWVLAAGVLAANAAC 1674
 Qy 1679 LATGVCVCIIRLHINORAVAVADKEVLEAPDEMEBCASRALIIEGORIAEMLSKIQG 1738
 Db 1675 LATGVCVCIIRLHINORAVAVADKEVLEAPDEMEBCASRALIIEGORIAEMLSKIQG 1734
 Qy 1739 LLOQASKOAOQIOPVQASWPKYEQFWAKGMNFIISGIQYLAGLSTLPGNPVAVASMAFS 1798
 Db 1735 LLOQASKOAOQIOPVQASWPKYEQFWAKGMNFIISGIQYLAGLSTLPGNPVAVASMAFS 1794

Qy 1799 AALTSPLSTSTTILINILGMLASQIAPAGATGVVSGLVCAAVSGIGLKVLDIILAG 1858
 Db 1795 ASITSPLTQNTLNLNIGWVAQIAPPASASAFVAGIACALIGLKVLDIILAG 1854
 Qy 1859 YGAGISGALVAFKINSGERKPSMEDVNNLLPGLISPGALVGVICAIIIRRHVPBGAVO 1918
 Db 1855 YGAGISGALVAFKINSGERKPSMEDVNNLLPGLISPGALVGVICAIIIRRHVPBGAVO 1914
 Qy 1919 WNRRLIAASRGNHAPTHYVTESSASORVYQOLGSLITTSILRRLHMTEDCEIPCCG 1978
 Db 1915 WNRRLIAASRGNHAPTHYVTESSASORVYQOLGSLITTSILRRLHMTEDCEIPCCG 1974
 Qy 1979 SWLRDPMWVCITLDFKXWILTSKLFPMMPGLPFVSCQKGYKGVAGTGIMTRPCCGAN 2038
 Db 1975 SWLRDPMWVCITLDFKXWILTSKLFPMMPGLPFVSCQKGYKGVAGTGIMTRPCCGAN 2034
 Qy 2039 ISGNVRLGSMRITGPKTCMNITWGTFPINCYTEGOCVEKPAWPFKVAIMRWVASEYAEVT 2098
 Db 2035 ISGNVRLGSMRITGPKTCMNITWGTFPINCYTEGOCVEKPAWPFKVAIMRWVASEYAEVT 2094
 Qy 2099 QHGSYHYITGLTTDNLKVPQOLPSPPEFSSWDVQOIHPPAPTPKPFPRFDEVSCVGLNSF 2158
 Db 2095 QHGSYHYITGLTTDNLKVPQOLPSPPEFSSWDVQOIHPPAPTPKPFPRFDEVSCVGLNSF 2154
 Qy 2159 VVGSQLPDCEPDPDVLMSMLTDPESHITAEFAARLARGSPPESSASSASQSLAPSLRAT 2218
 Db 2155 VVGSQLPDCEPDPDVLMSMLTDPESHITAEFAARLARGSPPESSASSASQSLAPSLRAT 2214
 Qy 2219 CTHGKADVMDVMDANLF---MGDDVTRIESGSKVYVLDLSDPMBEERSDLEPSIPSEY 2274
 Db 2215 CTHGKADVMDVMDANLF---MGDDVTRIESGSKVYVLDLSDPMBEERSDLEPSIPSEY 2274
 Qy 2275 MLPKRPPPALPAMARPDYNNPPLVESWKRPDYQATVAGCALPPPKPTPTPPRRRTYV 2334
 Db 2275 MLPKRPPPALPAMARPDYNNPPLVESWKRPDYQATVAGCALPPPKPTPTPPRRRTYV 2334
 Qy 2335 LSEBDSIGALQOALIKSRGQPPSGDSGLSGAGAGDSQOT-PPDELA-----LSENGS 2388
 Db 2335 LSEBDSIGALQOALIKSRGQPPSGDSGLSGAGAGDSQOT-PPDELA-----LSENGS 2384
 Qy 2389 ISSMPLEGEIADPDLEPEQVEPOPPOGVAAPSGDSGSMSTCS-EEDDSVVCCSMYS 2447
 Db 2385 ISSMPLEGEIADPDLEPEQVEPOPPOGVAAPSGDSGSMSTCS-EEDDSVVCCSMYS 2424
 Qy 2448 WTGALITPCSPBEEXLPINPLSNSILRYHNKYCTTYSASLRAKVTIDRMQVLDIYYD 2507
 Db 2445 WTGALITPCSPBEEXLPINPLSNSILRYHNKYCTTYSASLRAKVTIDRMQVLDIYYD 2484
 Qy 2508 SVLKDILKLAASKYARALLTMEBAQOLTPHSGARSKYRGAGAEVNSLSGRAVNIKSWKD 2567
 Db 2485 SVLKDILKLAASKYARALLTMEBAQOLTPHSGARSKYRGAGAEVNSLSGRAVNIKSWKD 2544
 Qy 2568 LLEDESEPIPTTIMAKNEVFCVDPPTKGGKKAARLVYDLAGRVCEKMAALYDITQKLPQA 2627
 Db 2545 LLEDESEPIPTTIMAKNEVFCVDPPTKGGKKAARLVYDLAGRVCEKMAALYDITQKLPQA 2604
 Qy 2628 VMGASGYFOYSPAOVRFEFLAAMAEKKDPMGPFSDYTRCFDSTVTERDIRTEBSIYRACSL 2687
 Db 2605 VMGASGYFOYSPAOVRFEFLAAMAEKKDPMGPFSDYTRCFDSTVTERDIRTEBSIYRACSL 2664
 Qy 2668 PBEAHTAHSILTELYVGGPMFNKSGQTCYRRCASGVLTSMQNTTITCYKALAAKCA 2747
 Db 2665 PBEAHTAHSILTELYVGGPMFNKSGQTCYRRCASGVLTSMQNTTITCYKALAAKCA 2724
 Qy 2748 AGIAPTMLVCGDDLVVISESGTEEDERNRLAFTFATRYASAPGDPPEPEYDELITLS 2807
 Db 2725 AKLOQCTMLVNGDDLVVISESGTEEDERNRLAFTFATRYASAPGDPPEPEYDELITLS 2784
 Qy 2808 CSSNVSVAGPOGRRRYLTDPPTPIARAAMEVTRHSPVNSWGNITQYAPTIYARVLT 2867
 Db 2785 CSSNVSVAGPOGRRRYLTDPPTPIARAAMEVTRHSPVNSWGNITQYAPTIYARVLT 2844

QY 2868 MTHFESILMAQDITLDONLFEMYGAVYSPLDLPALIERLHGLDAFSLHTYTHPELTRY 2927
DB 2845 MTHFESILMAQDITLDONLFEMYGAVYSPLDLPALIERLHGLDAFSLHTYTHPELTRY 2904
QY 2928 ASALRLKGLAPPLPAWKSRAAARASLISGCGRAAVCGRYLFMMAVTXYLKTPLPLPARLL 2987
DB 2905 ASALRLKGLAPPLPAWKSRAAARASLISGCGRAAVCGRYLFMMAVTXYLKTPLPLPARLL 2964
QY 2988 DLSWFTVAGAGGDIYHSVSRARPRLLFGLLLFYVGLFLPLPAR 3033
DB 2965 DLSWFTVAGAGGDIYHSVSRARPRLLFGLLLFYVGLFLPLPAR 3010
RESULT 6
GNMVC3
genome polypeptide - hepatitis C virus (strain HCV-1)
N:contains: capsid protein C; envelope protein M; hepacivirin (HC 3.4.21.98) (nonstruct
protein NS4; nonstructural protein NS5b; nonstructural protein NS5
C;Species: hepatitis C virus
C/Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C/Accession: A39166; PQ0403; PQ0404
R/Chao, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Col
Proc. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991
A>Title: Genetic organization and diversity of the hepatitis C virus.
A/Reference number: A39166; PMID:91172826; PMID:1848704
A/Accession: A39166
A/Molecule type: mRNA
A/Residues: 1-3011 <CHO>
A/Cross-references: UNIPROT:P26664; GB:M62321; NID:g329873; PDB:AAA45676.1; PID:g329874
R/Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, F.L
J. Gen. Virol. 73, 1131-1141, 1992
A>Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to e
A/Reference number: PQ0393; PMID:92268871; PMID:1316939
A/Accession: PQ0403
A/Molecule type: genomic RNA
A/Residues: 1577-1633 <CHA>
A/Cross-references: DDBJ:D10128
A/Experimental source: Isolates E-b16
A/Accession: PQ0404
A/Status: preliminary
A/Molecule type: genomic RNA
A/Residues: 1577-1633 <CH2>
A/Experimental source: Isolates E-b17
C/Superfamily: hepatitis C virus genome polypeptide
C/Keywords: ATP; capsid protein C; envelope protein; glycoprotein; hydrolase; nonstructura
F:1-115/Product: capsid protein C #status predicted <CPC>
F:115-191/Product: envelope protein M #status predicted <EMP>
F:192-389/Product: major envelope protein E #status predicted <MEB>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepacivirin #status predicted <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEH motif
F:1616-1862/Product: nonstructural protein NS4 #status predicted <NS4>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <NS4b>
F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,305,417,423,430,448,476,540,556,576,623,645,1213,1255,2041,2077,22

DB 121 KVIDTLTCGADLMGTYIPVVGAPLGGAAALAHGVLEEDGVNATGNLPGCSFSIFLLA 180
QY 181 LLSCTTTPVSAEAVNISTG-YMTNDCTNDISTMQLAAVLHAYGVCPCCKVGNASQW 239
DB 181 LLSCTTTPVSAEAVNISTG-YMTNDCTNDISTMQLAAVLHAYGVCPCCKVGNASQW 239
QY 240 IPVSNVAVQPPAL-TQGLRTH DMVMSATLCSALYVDLCGGVMAAQMFIVSPQH 297
DB 240 VAMTPTVAT-RDGKLPAQ-LRRHIDLLVGSATLCSALYVDLCGGVMAAQMFIVSPQH 297
QY 298 HMFVDDCNCSITPGLTHBAMMDMMMSPTATMILAAMRVEVITDIISGHWGTMF 357
DB 298 HMTTQGCNCSITPGLTHBAMMDMMMSPTATMILAAMRVEVITDIISGHWGTMF 357
QY 358 GLAEPMSOGAAKVVVILLAAVADARTHTVGSAAQOTGTLTSLFDMGPRQKQLVNTN 417
DB 358 GLAEPMSOGAAKVVVILLAAVADARTHTVGSAAQOTGTLTSLFDMGPRQKQLVNTN 417
QY 418 GSWHINRTALNCNDSLHTGFLASLEFYTHSFSNSSGCPERMASCRSTIEAFRVGMAQYEDN 477
DB 418 GSWHINRTALNCNDSLHTGFLASLEFYTHSFSNSSGCPERMASCRSTIEAFRVGMAQYEDN 476
QY 478 VTNPEDMRPGCMHPPROCGVSAKTVGCGPYVCTPSPVVVGTDRLGAPTYTGCEMTD 537
DB 478 GSGP-DQRPYCMHPPKPCGIVPAKSVCGPYCTPSPVVVGTDRLGAPTYTGCEMTD 535
QY 538 VFLNSTPRLPGSGMFGCTMNNSSGTYTKTCGAPPCRTADFNASTDLCTPDCFRKHPTT 597
DB 538 VFLNSTPRLPGSGMFGCTMNNSSGTYTKTCGAPPCRTADFNASTDLCTPDCFRKHPTT 593
QY 598 YLKCGSGFWLTPRCLIDYPRLMHYPCVTNYTIFKIMRYGVGEHRLTAACNFTRGDRCN 657
DB 598 YSRCSGSGWITPRCLIDYPRLMHYPCVTNYTIFKIMRYGVGEHRLTAACNFTRGDRCD 653
QY 658 LEDRDSLSPLSTTEMALTPCSYSDLPALSTGLLHONIVDQVCMVGSALATKYI 717
DB 658 LEDRDSLSPLSTTEMALTPCSYSDLPALSTGLLHONIVDQVCMVGSALATKYI 713
QY 718 VRMEWVILLFLLADARVACIAMLILGQEAALAEKLVIIHAASASCSNGLFYVIFV 777
DB 718 VRMEWVILLFLLADARVACIAMLILGQEAALAEKLVIIHAASASCSNGLFYVIFV 773
QY 778 AAWYIKGVPLATYSLTGLMSFSLILLALPQAAVADASYHGOI GAALLVMITLFTLP 837
DB 778 AAWYIKGVPLATYSLTGLMSFSLILLALPQAAVADASYHGOI GAALLVMITLFTLP 833
QY 838 CYKTLISFLMMLCYLLLGEMAOEMAPRMQVGGRCGIIWAVALPFGVFPITKLL 897
DB 838 CYKTLISFLMMLCYLLLGEMAOEMAPRMQVGGRCGIIWAVALPFGVFPITKLL 893
QY 894 YKRYTISCLMWLOFLTRVBAQLHVMIPLVNVRGDAVILLCAVHPVLVFDITKLL 953
DB 894 YKRYTISCLMWLOFLTRVBAQLHVMIPLVNVRGDAVILLCAVHPVLVFDITKLL 953
QY 958 TPMSMAASGLRDLAAVEPIIFSPMEKVIYWGAEYTAACDILHGLPVSARLREVLIG 1017
DB 958 TPMSMAASGLRDLAAVEPIIFSPMEKVIYWGAEYTAACDILHGLPVSARLREVLIG 1013
QY 1018 PADGTSGMGSLAPITAYAAOTRGLCTIVMSMGRPKTQDAEIOVLSTVTSPLCTS 1077
DB 1018 PADGTSGMGSLAPITAYAAOTRGLCTIVMSMGRPKTQDAEIOVLSTVTSPLCTS 1073
QY 1078 ISGVLMYVGAAGNKTLAGSGPQYOMWSAEGDLVGPSPGRTKSLPCTCGAVDLYV 1137
DB 1078 ISGVLMYVGAAGNKTLAGSGPQYOMWSAEGDLVGPSPGRTKSLPCTCGAVDLYV 1133
QY 1138 TRNADVIPARRGDKRGALSPRLSTYKSSGGGVLCPRHAYGVFPAAVCSGVAKSI 1197
DB 1138 TRNADVIPARRGDKRGALSPRLSTYKSSGGGVLCPRHAYGVFPAAVCSGVAKSI 1193
QY 1198 DFIPEVETDIYTRSPSTGSDTPPAVPTOVGVLAHAPTSKSKTKVVAAYAGYKLV 1257
DB 1198 DFIPEVETDIYTRSPSTGSDTPPAVPTOVGVLAHAPTSKSKTKVVAAYAGYKLV 1253

1258 LNPVAAATLGEAYLSKAGINPNIRGTGRTVTTGAPITVSTYSGFLADGGCAGAYDII 1317
 1254 LNPVAAATLGEAYLSKAGINPNIRGTGRTVTTGAPITVSTYSGFLADGGCAGAYDII 1313
 1318 ICDECHAVDSTTTIGITGLVLDQAEATGAVLTVLATATPPGVSVTPHPNIEEVALQOEGBI 1377
 1314 ICDECHAVDSTTTIGITGLVLDQAEATGAVLTVLATATPPGVSVTPHPNIEEVALSTGBI 1373
 1378 PFYGAATLSTYIKGGRHILFCHSKKKCBELAAALBGMKINSVAYYRGIDVSTYIPQGDVY 1437
 1374 PFYGAATLSTYIKGGRHILFCHSKKKCBELAAALBGMKINSVAYYRGIDVSTYIPQGDVY 1433
 1438 VVATATMTGXTGDPDSVIDCNVAVTQVYVDFSLDPTFTITQIVODAVSRQGRGTGR 1497
 1434 VVATATMTGXTGDPDSVIDCNVAVTQVYVDFSLDPTFTITQIVODAVSRQGRGTGR 1493
 1498 GRLGIRYVSTGERASGMFDSVVLCECYDAGAAWYELTPESETTVLRAYFNTPGLPVQCD 1557
 1494 GRLGIRYVSTGERASGMFDSVVLCECYDAGAAWYELTPESETTVLRAYFNTPGLPVQCD 1553
 1558 HLEFHEAVFTGLTHIDAHFLSOTKSGENFALTVYQATVCARAAQAPPSMDVMMKCLTR 1617
 1554 HLEFHEAVFTGLTHIDAHFLSOTKSGENFALTVYQATVCARAAQAPPSMDVMMKCLTR 1613
 1618 LKPTLVGPTPLLYRSGVTNEVTLTHPTKYIATCMQADLEWMTSTWYLAGVLAAYVAY 1677
 1614 LKPTLVGPTPLLYRSGVTNEVTLTHPTKYIATCMQADLEWMTSTWYLAGVLAAYVAY 1673
 1678 CLATGCVCIIGRLHINQRAVVAADKVELYEADEMEECASRAALIEEGQRIAEMLKSKTIQ 1737
 1674 CLATGCVCIIGRLHINQRAVVAADKVELYEADEMEECASRAALIEEGQRIAEMLKSKTIQ 1733
 1738 GLTQASQADODIOFTVOASMPKVEQPAKXHMNFIISGIOYIAGSTLPGNPAVASMAY 1797
 1734 GLTQASQADODIOFTVOASMPKVEQPAKXHMNFIISGIOYIAGSTLPGNPAVASMAY 1793
 1798 SAALTSPLSTTTILNLTGMLASQIAPAGATGFVSVGLVGAAGVSGIGKLVVDILA 1857
 1794 SAALTSPLSTTTILNLTGMLASQIAPAGATGFVSVGLVGAAGVSGIGKLVVDILA 1853
 1858 GYGAGISGALVAFKIMSEKPSMEDVNLPGILSPGALVGVICAAILRRHVGBEGAV 1917
 1854 GYGAGISGALVAFKIMSEKPSMEDVNLPGILSPGALVGVICAAILRRHVGBEGAV 1913
 1918 QMANRLIAPASRGHVAFTHYTESDAQRTQLIGSTLTSRLRLHNMITEDECPICG 1977
 1914 QMANRLIAPASRGHVAFTHYTESDAQRTQLIGSTLTSRLRLHNMITEDECPICG 1973
 1978 GSWLRDWDVWCTITLDFKNNLTSLPKMPGLPVSOGKKGWAGGIMTTRCPGCA 2037
 1974 GSWLRDWDVWCTITLDFKNNLTSLPKMPGLPVSOGKKGWAGGIMTTRCPGCA 2033
 2038 NISGNVRLSGNRITGPKTCMNIWQCTFPINCYTEGQCVKPAKPAKVAIWRVAASEVAV 2097
 2034 NISGNVRLSGNRITGPKTCMNIWQCTFPINCYTEGQCVKPAKPAKVAIWRVAASEVAV 2093
 2098 TQHSYHITIGITDNLKVPQCLPSPERFSWVDGQIHRFAPTPKPFRRDEVSFCVGLNS 2157
 2094 TQHSYHITIGITDNLKVPQCLPSPERFSWVDGQIHRFAPTPKPFRRDEVSFCVGLNS 2153
 2158 FVVSQALCPDEPDPDLVMSMLTDPSSHITAEARLARGSPSSASSASQSLASPSIRA 2217
 2154 FVVSQALCPDEPDPDLVMSMLTDPSSHITAEARLARGSPSSASSASQSLASPSIRA 2213
 2218 TCTTHGKAYDVMDVANDLFG---WGDVTRIESSGKVVVLDLDMVEERSDLEBSIPSE 2273
 2214 TCTTHGKAYDVMDVANDLFG---WGDVTRIESSGKVVVLDLDMVEERSDLEBSIPSE 2273
 2274 YMLPKRRPRLPAMARPDYNPPLVESMKRPYQATYAGCLLPPRPKPTPPRRRRRTV 2333
 2274 ILKRSRRAQALPVWARDYNPPLVETWKKDPYEPVVHGCPLPPKSPDPVPPPKKRTV 2333

2334 GLESDSICDALQOLAIKSGFQPPPSGDSGLSTGAGAADSGQTPEDEALSETGSISSNP 2393
 2334 VLTSTSTLALAEALATRSFGSSSTSGITGDNNTTSSSEAPSGCPSD---SDAESVSSMP 2389
 2394 PLEGEALDPDUEPEOVPEQPPQGGVAAAPGSDSGSWSTCSEF--DSDVCCMSYSWTGA 2451
 2390 PLEGEALDPDUEPEOVPEQPPQGGVAAAPGSDSGSWSTCSEF--DSDVCCMSYSWTGA 2429
 2452 LITPCSPBEKLPINPLNSLIRYANKYCTTTKASIRAKVVTDRMOVLDSYDSVYK 2511
 2430 LITPCSPBEKLPINPLNSLIRYANKYCTTTKASIRAKVVTDRMOVLDSYDSVYK 2489
 2512 DIKLAASKYATRLITMEBACOLTPPHSASKYGFKAKEVRSLSGRAVHNIKSVMDLDED 2571
 2490 EYKAAASKYKNTLSVEBACOLTPPHSASKYGFKAKEVRSLSGRAVHNIKSVMDLDED 2549
 2572 SETPIPTTIMAKNEVEFCVPTKGGKKAARLIYVDPDLYRVCEKMAIYDITOKLPOAVMGA 2631
 2550 SETPIPTTIMAKNEVEFCVPTKGGKKAARLIYVDPDLYRVCEKMAIYDITOKLPOAVMGA 2609
 2632 SYGFOYSPOKVEFLVQAMKSKTTPMGFSYDRCPDSTVETSDIKTEEALTYOCCDLPQA 2691
 2610 SYGFOYSPOKVEFLVQAMKSKTTPMGFSYDRCPDSTVETSDIKTEEALTYOCCDLPQA 2669
 2692 HTAHSILTERLYVGGPMFNSKQOTCYRRRCRASGVLTTSMGNTTCYVKAALACRAAGII 2751
 2670 HTAHSILTERLYVGGPMFNSKQOTCYRRRCRASGVLTTSMGNTTCYVKAALACRAAGII 2729
 2752 APTMLVCGDDLVYISGCTEBEDERNLAFBAMTRYSAAPPCDDPRPEYDLELTSCSN 2811
 2730 APTMLVCGDDLVYISGCTEBEDERNLAFBAMTRYSAAPPCDDPRPEYDLELTSCSN 2789
 2812 VSVVLPQGRRRYVLTPTPTPIARAAMEYTRHSPVNSMNGNIIOYAPTIVARVLTMTMF 2871
 2790 VSVVLPQGRRRYVLTPTPTPTPIARAAMEYTRHSPVNSMNGNIIOYAPTIVARVLTMTMF 2849
 2872 FSIILMAQDTPDQNTFENYGAIVSVSPLDPAIETRLGLDAFSLHTTYTPHELTREVASL 2931
 2850 FSIILMAQDTPDQNTFENYGAIVSVSPLDPAIETRLGLDAFSLHTTYTPHELTREVASL 2909
 2932 RKLGAAPLPAWKRARAVRASLISRGRAVCGRYLFEMAYVTKLKLTPLEARLIDISS 2991
 2910 RKLGAAPLPAWKRARAVRASLISRGRAVCGRYLFEMAYVTKLKLTPLEARLIDISS 2969
 2992 WFTYAGAGGDIYHSVSRARPLLLFGLLTFVGVCLFLLPPAR 3033
 2970 WFTYAGAGGDIYHSVSRARPLLLFGLLTFVGVCLFLLPPAR 3011

RESULT 7
 genome polypeptide - hepatitis C virus (strain Taiwan)
 N:Contains: capsid protein C; envelope protein M; hepatitis virus (EC 3.4.21.98) (nonstructural protein NS4b; nonstructural protein NS5)
 C:Species: hepatitis C virus
 A:Note: host Homo sapiens (man)
 C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
 C:Accession: A40244
 R:Chen, P.-J.; Lin, M.-H.; Tai, K.-F.; Liu, P.-C.; Lin, C.-J.; Chen, D.-S.
 Virology 188, 102-113, 1992
 A:Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the
 A:Reference number: A40244; MUID:92230206; PMID:1314449
 A:Accession: A40244
 A:Molecule type: genomic RNA
 A:Releases: 1-3010 <CHE>
 A:Cross-references: UNIPROT:P29846; GB:M84754
 C:Superfamily: hepatitis C virus genome polypeptide
 C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural
 F.1-115/Product: capsid protein C #status predicted <CPC>
 F.116-191/Product: envelope protein M #status predicted
 F.192-389/Product: major envelope protein E #status predicted <ME>
 F.390-729/Product: nonstructural protein NS1 #status predicted <NS1>
 F.730-1006/Product: nonstructural protein NS2 #status predicted <NS2>

F:1007-1615/Product: hepacivirin #status predicted <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <NA>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <NA>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS>
F:196,209,233,234,250,305,325,417,423,430,448,552,540,556,576,623,645,1213,1255,2041,207

Query Match 73.9%; Score 11911.5; DB 1; Length 3010;

Best Local Similarity 71.1%; Pred. No. 0;

Matches 2163; Conservative 352; Mismatches 484; Indels 45; Gaps 8;

```
QY 1 MSTNPKRQRTKNTNRNRPQDVKFGSGQIVGCVYLLPRGRPLGYRATRKTSERSQPRG 60
Db 1 MSTNGKPRQRTKNTNRNRPQDVKFGSGQIVGCVYLLPRGRPLGYRATRKTWERSQPRG 60

QY 61 RRPPIKDRSRSTKSGKRGKPYWPLYGNBGLGAGMLSPRGRSPWGNDRPHNSRNVG 120
Db 61 RRPPIKARQPEGRMAQPGYPMPLGNBGLGAGMLVSPRGRSPWGMGPTDPRRRSRNLG 120

QY 121 KVIDTLTCGFADMGYIPVVGAPLGGVAPALAHGVLEBDGVNFAFGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADMGYIPVVGAPLGGVAPALAHGVLEBDGVNFAFGNLPGCSFSIFLLA 180

QY 181 LLSCTITTPVSAAEVKNISTGYMTNDCTNDSITWQLQAAVLHVGCVPCEKYNASQCM 240
Db 181 LLSCLTIPASAEVHNVSIGYHTNDCNSIYEEAADMIMHPCGVCPCVRENNSRCNV 240

QY 241 PVSNNVAPRGALTOGLRTHIDMVMSATLGSALIVGDLCSGVMLAQMFTVSPQHMF 300
Db 241 ALPTPLAARNNSVPTATIRRHVDLVGAALFCSAMVVGDLCSGVFVLSOLFTEFSPRRYET 300

QY 301 VQDNCSTIYPGTITGHRMADMMNNSPTATMLAAMRVEPTIDILIGAHGVNFGIA 360
Db 301 VQDNCSTIYPGHVGHRMADMMNNSPTTALVSQLRTPQAVVMVGAGHGVLAGIA 360

QY 361 YFSMOGAMAKVYVILLAAGVDAETHVGSAAQTGRLLSLFDMGPROKIQLVNTNGSM 420
Db 361 YFSMVGAMAKVLLVLLFAGVDSTIVSGTVARTHSLASLFTQASQKIQLVNTNGSM 420

QY 421 HIRKTLNCNDSLHTGFASLFTYTHFNSSGCEPERSACRSIEAPVNGALQYEDNVTN 480
Db 421 HIRKTLNCNDSLQTFELASLFTYAHFNSSGCEPERSACRSIDKFOGCGPTLYTE-- -48

QY 481 PEDMRPVCWHYPRGCGVSAKTVCGPVYCFTPSPVVGTTDLGAPTYTWGENEDVPL 540
Db 479 IQDPRFCWHYAPRPGCIIVPASQVCGPVYCFTPSPVVGTTDLFGAPITSMGENEDVLI 538

QY 541 LNSTRPPLSGWFCCTWNNSSGYTKTCGAPRCRTRADFNASTDLLCPTDCFRKHPTTYLK 600
Db 539 LNSTRPPOGWNFCCTWNNSTGFTKCGGPRCNI GGGGN-- NTLVCPDTCFRKHPEATYRK 596

QY 601 CGSGPMULTPRCLIDYPRRLMHPCTYNTYTFKIRMYVGVGHEHLLTAACNFTGRDCLND 660
Db 597 CGSGPMULTPRCVMVDYPRRLMHPCTYNTYTFKIRMYVGVGHEHLLTAACNFTGRDCLND 656

QY 661 RDSQSPLSLHSTTEWAILPCSYSDLPALSTGLLHONVVDQVQYHGSAPLTKXYVM 720
Db 657 RDSSELSPLSLHSTTEWAILPCSYSDLPALSTGLLHONVVDQVYHGSAAVSPALTK 716

QY 721 EWWILLFLLIADARVCACTMMLILGQAEALAEKYLILHAASNASGNGELYPVIFVAVM 780
Db 717 EYILLFLLIADARVCACTMMLILGQAEALAEKYLILHAASNASVAGHGTLSLFLVFCAAM 776

QY 781 YIYGRVVPPLATYSLTGLMSFSLILALPQQAIVYDASVHGQICALLVMTLFTLTPGX 840
Db 777 YIYGRVLPGAAYVALYGMPLLLILALPRAVYAMDEEMASCCGAAPVGVVLTLSLPHYK 836

QY 841 TLKSRFLMWLCYLLTIGEAMVQEMAPPMOYRGRDGIIMAVALFYGVGVFDITKMLLAVL 900
Db 837 MFLARLIMLOVYFTITRAEHLQVMIPLLVNNGSDAILITLCAAYFELLFDITKMLLAIL 896
```

```
QY 901 GPAYLLKCALTRVEYFVRAHALRMCMTARHLAAGRYVQNALALGHWGTGYIYDHLTPM 960
Db 897 GPLWVLQAGLIRIYFVPAQOGLIRACMLVRAAGHYVQNALMLKLTGYVVDHLTPL 956

QY 961 SDMAASGLRDLAVALVEPIESPMKRYIWMGAETPAAGDILHGLPVSARLREVLLEPAD 1020
Db 957 QDMWHTGRLDAVALVEPVFSDMETKIIYTWADPAACGDIIILGIPVSRRRRELLLEPAD 1016

QY 1021 GYTSKMSLLAPITRVAQOTRGLLGTIVSWMTGDKTEQAGEIOLVSTVOSPIGTSISG 1080
Db 1017 SLEGRMRLLPITRVAQOTRGLLGTIVSTLTDGDKQVBESEVQVSTVATOSPLATCING 1076

QY 1081 VLMVTYHAGNKTLIAGSRGVYQVYSSAEGDLVGPSPPTGKSLPECTCGADVLYVTRN 1140
Db 1077 VCMVTYHAGSKTLIAGSRGVYQVYSSAEGDLVGPSPPTGKSLPECTCGADVLYVTRH 1136

QY 1141 ADVTPARRGDKRGALLSPRLSTLTKSSGGPVLCPRGHANGVRAAVCSGVAKSIDFI 1200
Db 1137 ADVTPARRGDKRSRSLSPRLSTLTKSSGGPVLCPSGHVVGITPRAAVCTRGVAKADV 1196

QY 1201 PVETLIDVTRSPFTSDNSTPRAVPQYQVGYLHAPTSKSTKYVPAAYAQGYKVLVNP 1260
Db 1197 PVESMETMRSPVFTSDNSTPRAVPQYQVGYLHAPTSKSTKYVPAAYAQGYKVLVNP 1256

QY 1261 SVAATLFGAYLSKAHGINPRTGRTVTTGABITYSTYKSTFLADGGCAGAAVDIICD 1320
Db 1257 SVAATLFGAYLSKAHGINPRTGRTVTTGABITYSTYKSTFLADGGCAGAAVDIIMCD 1316

QY 1321 ECHAVDSITTIIGITVLDQAEATAGRLVLTATPRTGSAVTTPHNIEVVALGQGEIPEY 1380
Db 1317 ECHAVDSITTIIGITVLDQAEATAGRLVLTATPRTGSAVTTPHNIEVVALGQGEIPEY 1376

QY 1381 GRAIPLSYIKGGRHLIFCHSKKKCCDELAALRGKMSVAYYRGDLSVITPOGDVVVVA 1440
Db 1377 GKALPIETIKGGRHLIFCHSKKKCCDELAALRGKMSVAYYRGDLSVITPOGDVVVVA 1436

QY 1441 TDLMTGYTGDPSVIDCNAVTOYVDSIDPTFTITTOIVPODAVSHSRGRGTGRRL 1500
Db 1437 TDLMTGYTGDPSVIDCNAVTOYVDSIDPTFTITTOIVPODAVSHSRGRGTGRRL 1496

QY 1501 GIYRVYSTGERASGMFDSVULCECYDAGAAVYELTSPSTTYRLAYFNTPLPVCQDHL 1560
Db 1497 GIYRVYSTGERASGMFDSVULCECYDAGAAVYELTSPSTTYRLAYFNTPLPVCQDHL 1556

QY 1561 FWEAVFTGLTHIDHAFISQTKOSGENFAYLTAAYQATYCARAKAPPSPSDVMKKCLTRLP 1620
Db 1557 FWEAVFTGLTHIDHAFISQTKOSGENFAYLTAAYQATYCARAKAPPSPSDVMKKCLTRLP 1616

QY 1621 TLVGPPLLYRIGSVYNEVTLTHPVTKYIATCQADLEVMSTVWLAGVLAAVAAYCLA 1680
Db 1617 TLHGTPPLLYRIGSVYNEVTLTHPVTKYIATCQADLEVMSTVWLAGVLAALAAYCLT 1676

QY 1681 TGCVCILIRLHINORAVAPDKVLYEAFDMEERCASAAALIEGORTAEMLKSKIGQL 1740
Db 1677 TGSVAVIRIILSGPAPVAPPREVLYOEFDEMEERCASAAALIEGORTAEMLKSKIGQL 1736

QY 1741 COASROADIOPTVOASWPKYEOFWAKKMMNFISGIQYLAIGLSTPCGNPAVAMMAFSA 1800
Db 1737 QATYQOAAAPAVESKRTIEAFKANDMMNFISGIQYLAIGLSTPCGNPAVAMMAFSA 1796

QY 1801 LTPSLISTITLILNLIGWMLASQIAPPAAGATGVVSGLVGAAGSILGKLVLDILAGY 1860
Db 1797 ITSPITLSTITLILNLIGWMLASQIAPPAAGATGVVSGLVGAAGSILGKLVLDILAGY 1856

QY 1861 AGISGALVAFKIMSEKSMEDVNNLPGIISPGALVGVCAALIRRHVPBGEGAVQMM 1920
Db 1857 AGVAGALVAFKIMSEKSMEDVNNLPGIISPGALVGVCAALIRRHVPBGEGAVQMM 1916

QY 1921 NRIIFASRGHNVAPATHVTESSDAQRYTQGLISLTISLRRLHNMVTTEDCPIPCGSS 1980
Db 1917 NRIIFASRGHNVSTHVPESDAARVYTOQLISLTISLRRLHNMVTTEDCPIPCGSS 1976

QY 1981 LRDVMDWCTIITDCKMWTSKLFPKMPGLPFVSCQKGYKGVWAGTGIMTTRCPGANIS 2040
```

```

Db      1977 LRDVDMCTCTVLADPKTWLQSKLRLPGVPEFSCQRGVKGVWRSDGIMQTTCPGAGLDT 2036
Qy      2041 GNVRLGSRIRITGPTCNAINIQCPTINCYSQCQPKAPNFKAIWVAASEVAVTQH 2100
Db      2037 GHVKGSGMRIRIGPKTCSNTWHTGPTINAYTTGPTCPSPAPNVSRLMVAEEVEVRV 2096
Qy      2101 GSHYITGLTLDNLKVPCLPSPPEFVWDGQVIRPAPTPPPPRRDVSPCVGNISPV 2160
Db      2097 GDFHVTMTTNDVACPCQVPAPEFTEVDGRLRKYAPACKPLREEVSPGVGNQYV 2156
Qy      2161 GSOLPCDEPDVLMMLTDPSSHITAARLARGSPPESSASSASQASPLRATCT 2220
Db      2157 GSQLPCEPEPVAVLTMLTDPSSHITAARLARGSPPESSASSASQASPLRATCT 2216
Qy      2221 THGRKAVDWDANLF-----MCDVTRIESSKVVVLDLPMVEERDLEPISSEWML 2276
Db      2217 TRHPPDADLIEANLMLRQEMGNITRVESEKVVILDSFDPLRAEEDEREVSVAETLR 2276
Qy      2277 PKKRPPLPAMARPDYNPPLVESKRPDYOPATVAGCALPPRRTPTPPRRRTVGIS 2336
Db      2277 KSRKFPPLPWAPRDPNPPLLEPKDDVYPVTHGCPPLPVKAPLP1PPRRKTVLT 2336
Qy      2337 EDSICDALQOLAISFCQPPSPSGSLSTGAGAADSGQT--PPDEL-----ALSETGIS 2390
Db      2337 ESTYSALAEIATKTFG-----SSSSAASGTAATAPDPDPSDDGAGSVESCS 2386
Qy      2391 SMPLEGLGDPDLEPEOVEPQPPQGVAAAPGSDSGMSWSTCEED--DSVCCSMYSWT 2449
Db      2387 SMPLEGLGDPDLEPEOVEPQPPQGVAAAPGSDSGMSWSTCEED--DSVCCSMYSWT 2426
Qy      2450 GALTIPGSEBEKLPINPLNSLRLYHNKYCTTKSASLRKKTTPRMOMVLDYSYOSV 2509
Db      2427 GALTIPCAEBSKLPINPLNSLRLYHNKYCTTKSASLRKKTTPRMOMVLDYSYOSV 2486
Qy      2510 LKDIKLAASKYARLLTMEACQLTPPHSASKEYFGAKAYSLSGRAVNHISKYWKDL 2569
Db      2487 LKEMAKASTYKAKLSVEBAKCLTPPHSASKEYFGAKAYSLSGRAVNHISKYWKDL 2546
Qy      2570 EDSERPITTTIMAKNEVFCVDPFTKGKKAARLLIYPDIGVRVCEMALYDITQKLPQAVM 2629
Db      2547 EDTERPITTTIMAKNEVFCVDPFTKGKKAARLLIYPDIGVRVCEMALYDITQKLPQAVM 2606
Qy      2630 GASVGFQVSPARVEFLKAAEKDPMGFSYDTCFPTSTYTERDIRREESTYRACSLPE 2689
Db      2607 GSSYGFQVSPARVEFLKAAEKDPMGFSYDTCFPTSTYTERDIRREESTYRACSLPE 2666
Qy      2690 EAHTAHSLTERLYVCGEMFNSKQTCGYRCRAGSVLTSMGNTITCYVKAALAAACKAAG 2749
Db      2667 EAHQAIRLTERLYVCGEMFNSKQTCGYRCRAGSVLTSMGNTITCYVKAALAAACKAAG 2726
Qy      2750 IIAFMLVCGDDLVISSQGTEDERNLRAFTAMTRYAPGDPPREPYDLELITSCS 2809
Db      2727 LQDCMTLVCGDDLVISSQGTEDERNLRAFTAMTRYAPGDPPREPYDLELITSCS 2786
Qy      2810 SNVSVALGPOGRRRYVLTDPPTIARAAMEVRSHPNSMGNIOVAPITMAWMLMT 2869
Db      2787 SNVSVAHDAAGRRYVLTDPPTIARAAMEVRSHPNSMGNIOVAPITMAWMLMT 2846
Qy      2870 HFFSILMAQDVLQDNLFEMYGAVSVSLDPAIIRLHGLDAFSLHTYTPHELTRVAS 2929
Db      2847 HFFSILMAQDVLQDNLFEMYGAVSVSLDPAIIRLHGLDAFSLHTYTPHELTRVAS 2906
Qy      2930 ALRKUAGPLPAMKSRAPAVRASLSIRGGAACVGRYVFNMAVYKTLKLTLPPEARLIDL 2989
Db      2907 CLRKGVGPPLPAMRHARSVRAKLDSOGGRAATCRRYVFNMAVYKTLKLTLPPEARLIDL 2966
Qy      2990 SSMFVVGAGGDIYHVSGRARPRLLLFGLLLFVVGFLIPAR 3033
Db      2967 SKMFAVAGYGDDIYHLSRPARPRMFLCLLLSVGVGVIYLLPNR 3010

```

```

S40770
genome polypeptide - hepatitis C virus
N:Contains: capsid protein C, envelope protein M, hepatitisin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
A:Accession: S40770; PC1285
R:Okamoto, H.
submitted to the EMBL Data Library, March 1992
A:Reference number: S40770
A:Accession: S40770
A:Molecule type: genomic RNA
A:Residues: 1-3011 <OK>
A:Cross-references: UNIPROT:003463; EMBL:D10749; NID:G221586; PID:BAA01582.1; PID:G2215
R:Okamoto, H., Okada, S., Sugiyama, Y., Yotsumoto, S., Tanaka, T., Yoshizawa, H., Tsuda,
Jpn. J. Exp. Med. 60, 167-177, 1990
A:Title: The 5'-terminal sequence of the hepatitis C virus genome.
A:Reference number: PC1284; MUID:91013116; PMID:2170712
A:Accession: PC1285
A:Molecule type: genomic RNA
A:Residues: 1-513 <OK>
A:Cross-references: GB:D00831; NID:G221511; PID:BAA00705.1; PID:G221512
A:Experimental source: isolate HC-01
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polypeptide; serin
F:2-115/Product: capsid protein C #status predicted <CC>
F:116-191/Product: envelope protein M #status predicted <EM>
F:192-389/Product: major envelope protein E #status predicted <ME>
F:730-1006/Product: nonstructural protein NS1 #status predicted <NS1>
F:1007-1615/Product: hepatitisin #status predicted <NS>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1332-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <NA>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <NB>
F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS>

Query Match 73.8%; Score 11895; DB 1; Length 3011;
Best Local Similarity 71.4%; Pred. No. 0;
Matches 2172; Conservative 349; Mismatches 481; Indels 40; Gaps 12;

Qy      1 MSTNPKPKRTKNTNRNPODVKFPFGQIYGVYLLPRRGRGLGVRATRTKTSRSPRG 60
Db      1 MSTNPKPKRTKNTNRNPODVKFPFGQIYGVYLLPRRGRGLGVRATRTKTSRSPRG 60
Qy      61 RRPIPKDRSTGKSGKRGYPMPLVYNGELGMAWMLSPGSRPSWGPNDPRHSRVNG 120
Db      61 RRPIPKDRSTGKSGKRGYPMPLVYNGELGMAWMLSPGSRPSWGPNDPRHSRVNG 120
Qy      121 KYIDLTGTFADLNGYITVWGAAPLGVARALAHGVRVLEQVNPATGMLPGCSFSIFLLA 180
Db      121 KYIDLTGTFADLNGYITVWGAAPLGVARALAHGVRVLEQVNPATGMLPGCSFSIFLLA 180
Qy      181 ILSCITVPSAAEVNISTG--YMTNDCTNSITWLOAAVLVHVGCVCEKRVNASOCM 239
Db      181 ILSCITVPSAAEVNISTG--YMTNDCTNSITWLOAAVLVHVGCVCEKRVNASOCM 239
Qy      240 VAMPTVAT--RDGKLPAQ--LRHIDLLVGSATLCSALYVGLDLSGVFLIGQLFTFSR 297
Db      240 VAMPTVAT--RDGKLPAQ--LRHIDLLVGSATLCSALYVGLDLSGVFLIGQLFTFSR 297
Qy      298 HMFVQDCNCSITPGITTHRAKMMMMWSPATMTIILYANRVPVILDIISGATWGMF 357
Db      298 HMTTQCNCSITPGITTHRAKMMMMWSPATMTIILYANRVPVILDIISGATWGMF 357
Qy      358 GLAFVSGMAKVVVILLAAVDARTHTVGSAAQTGRLTSLFDMGPRQKTLQVNTN 417
Db      358 GLAFVSGMAKVVVILLAAVDARTHTVGSAAQTGRLTSLFDMGPRQKTLQVNTN 417
Qy      418 GSHINIRNALCNDSLHTGTASTLFTYHSFNSSGCPERMSACRSIEAFVGMGLQYEDN 477
Db      418 GSHINIRNALCNDSLHTGTASTLFTYHSFNSSGCPERMSACRSIEAFVGMGLQYEDN 477

```

QY	478	UTNPEDMA PYCMHNP PROCGVNSAATVCGPVYCFPSPVVMVGTDRCLGAPRTYMGENTD	537
Db	477	GSBP -DQBPYCMHNP KPCKGIPVAPASVCGPVYCFPSPVVMVGTDRSGAPRTYMWKANDTD	535
QY	538	VFLNSTRPPLGSMWEGCTMNMNSGGYTKYCGAPPCRTRADFNASTDLCPTPCFRKHPPTT	597
Db	536	VFLNSTRPPLGNWGCCTMNMNSTGTXTYCGAPPCYIGGGN--NTHLCPTDCFRKHPAT	593
QY	558	YLKCGSGMWLTPRCLIDYPRIMHYPCIVNVTIFKIRMYVGVGEHRLTAACNFTRGDRCN	657
Db	554	YSRCGSGMWITPRCIVDPYRIMHPCTINVTIFKIRMYVGVGEHRLTAACNMTGGEBCD	653
QY	658	LEDRRSQDLSPLNHSSTEMALIPCGSYDLPLASTGLLHHONINIVOMYMGJSPALTXI	717
Db	654	LEDRRSLSPLNHSSTQOVPSCFTLLPALSTGLLHHONIVDVQVLYGVGJISAWA	713
QY	718	VRMWVILLFLLADARVACILMMLILLGOEAALKEKVLIIHAASPAACNGFLYFVFEV	777
Db	714	IKMEYVLLFLLADARVCSLMMMLILLSAEALENLVIIHAASLACTRGLVSPVLFEC	773
QY	778	AAWYIKGRVPLATSLTSLGMSFSLILLALPQAAAYASVHGQGAALLWITLTPTP	837
Db	774	FAWYIKGRVWPAAALYALGMWFLILLALLPORAALDTEVAASCGGVALLGMALTLSP	833
QY	838	GKTLTSLRFLMWLCVLLTLGEMAMVQEMAPMOWRGGRGIIIAWVAFPGVFPDITKML	897
Db	834	YKRCISWCLMWLOFLTRVEHQHVMVPLATNGRGRVAILMKVNHPTLVFDITKLL	893
QY	898	AVLGAVALLLKALTREPVFVRHAALLMCTMAHLAGRYYOMALLAGRWGTIYIDL	957
Db	894	AVLGLMTILQASLLKVYFVVRQGLRICALARKVGVGHVYOMALIKGALTGVVYHNL	953
QY	958	TPMSWMAASGLRDLAVANVEPIIFESMEKKVITYMGAEITAACGDILHGLEFVSARLREVLG	1017
Db	954	TPLRMANANGRLDLAVAEPPVFSQMETKLLTWGADITACGDIIINGLFSVSAKGEIILG	1013
QY	1018	PADGYTSKMSLILAEITAYAOOTRGLLGTIYVMSGRKTEQAGRIQVLSVTQSLFCTS	1077
Db	1014	PADGWSVSKMRLAETITAYAOOTRRLGCTITSLTGRDNQVGEHVQVSTVAQTLFATC	1073
QY	1078	ISGLMTVYHGAKNKTLAGSRGPVYOMYSAAEGDVGMPSPPTKSLEPCTCGADVILV	1137
Db	1074	INGVCMVYVHGAQRTTISASPKGPVQMTYNVDQDVGMPAPGARSILPTCGSSDILV	1133
QY	1138	TRNDVITARRRGDRGALLSRPLSTLKGSSGCVLCPRGHAGVFPFAACSRGVAKSI	1197
Db	1134	TRHADVITVRRRGDSRGLSLRPISLYKSGSGGFLCPAGHVNGIFPAAACTRVAACAV	1193
QY	1198	DFIPEVETLIDIVTRSPFSDNSTPPAPVOTYQVGYLHAAPTSGSKSTKVVAALAAQGYKLV	1257
Db	1194	DFIPEVSLETMTRSVFTDNSPPAPVPSFOVAILHAAPTSGSKSTKVVAALAAQGYKLV	1253
QY	1258	LNPSYAALTLGFGAYILSKAHGINPNIRTGVRVYTTGAPITYSTYGFKLADGCGAGADYII	1317
Db	1254	LNPSYAALTLGFGAYVMSKAHGIDPNIRTRGRTITGSPITYSYGFKFLADGCGSGADYII	1313
QY	1318	ICDECHAVDSFTLIGVTLQDAETAGVRLITVLAATPBGSTPHPHIEVALGQBEI	1377
Db	1314	ICDECHSDAISVLIGVTLQDAETAGRLVVLVLAATPBGSTIVPHANIEEVALSTGEI	1373
QY	1378	PFYGRALPLSYIKGGRHILFCHSKKKCBELAAALRGKMLNSAAYRGADVASIPTQGVV	1437
Db	1374	PFYGRALPLBEAIKGRHILFCHSKKKCBELAAKVALGVMNAVAYRGADVIVIPRSGVV	1433
QY	1438	VVATDALTMTGYTGDSDVIDCNVAATOVVDSLPDFTTITQIIVFQDASVSSQRRGRTGR	1497
Db	1434	VVATDALTMTGYTGDSDVIDCNTCTVTOVDFESLDPFTLETTITFQDASVSTRGGRTR	1493
QY	1498	GRIGIYRYSVGERBSGMPDSVVLCECDAGAAWELTPSETTVLRAVFMTPGLPYVOD	1557
Db	1494	GKPGIYRVAAPGERSGMPDSSIICECTDTCGAMETLPAETTVLRAVMMTPGLPYVOD	1553

QY	1558	HLEFMEVFPGGLTHIDANFLSOTKOSGEFALATYQATVCARAKAPPSNDVMKKSLTR	1617
Db	1554	HLEFMEVGFGLTHIDANFLSOTKOGGEFPLVAYQATVCARAKAPPSNDQMKCLIR	1613
QY	1618	LKPLVLGPPLLYRLGASVTNEVTLTPHVKIATCMQADLEVMTSTWYLAGVLAABAAY	1677
Db	1614	LKPLGLGPPLLYRLGAVOGEVTLTHPVKYIMTCSADLEVMTSTWYLVGGVLAALAAAY	1673
QY	1678	CLANGCVCIIGRIHINQAVVAPDXEVLVEAFDEMEBCASBAALIIEBORIAEMLKSKIQ	1737
Db	1674	CLSTGCVLVJGRVUSGRPAIIPDEVLREFDEMEBCQHLPYIEQMMILAEOFKQAL	1733
QY	1738	GLLQOASKOADIQPTVQASWPKVQOFNAKHMNNTISGIQYLAGLSTLPGNPAVASMAF	1797
Db	1734	GLLQTSARQGEVIAPTVOGTMQKLEAFNAKHMNNTISGIQYLAGLSTLPGNPAIASMAF	1793
QY	1798	SAALTPLSTSTIIILNIIIGWLAQIAPRAGATGFVUSGLVGAASIGLGTKLVIIILA	1857
Db	1794	TAAPTPLTTSQTLFNLIIIGWVAQAQIAPGATVFSVGLAGAASVSGIRLVIIILA	1853
QY	1858	GYGAGISGALVAFKISGEKPEMEDVNNLPILSPGALVGVCAALIRHNVPGEGAV	1917
Db	1854	GYGAGVAGALVAFKISGSLPSTEDLVNLPILSPGALVGVCAALIRHNVPGEGAV	1913
QY	1918	QMMNRLIAFASRGNHVAPTHYVTESDASORVTLIGSLITTSILRLHNMTEDCPLPG	1977
Db	1914	QMMNRLIAFASRGNHVSPHYVESDAARVAILSLTIVQLRHLHQMLSESTTPCS	1973
QY	1978	GSMRLDPVMPVCCILNDFKNWLTSLPRMPBLPVSCOKYKGVWAGTQIMTRCCGA	2037
Db	1974	GSMRLDMDCICVLSDFEYTKTLKTLMPBLPFIPIVSCOHGKGVWRDQIMHTRCCGA	2033
QY	2038	NISGNVRLGSMRTIGBTKCMINMOGTFPINCYBQCVPKAPNPKVAILMVAASEVAEV	2097
Db	2034	EIGHVANGMRLVGBTKCRANMMSGFPPINAVTTGPTPLPAPNVTFAIMVASEVEI	2093
QY	2098	TQHGSHYITGLTDMVKPQCLPSEPFNSWDGVIQIHRAPATKPFPRDVSCTVGLNS	2157
Db	2094	RRVGDFFHYVGMITMLKPCQVPSEPFTELDGRLHFRAPCCPFLIREVSFRVGLHD	2153
QY	2158	FVVGSQLPCPEPBDVTLSMLTDSHSHTAEPAARLRGSPSESASSAQLAPSHRA	2217
Db	2154	YVPGSQLPCEPBDVAVLNMLTDSHSHTAAAGRLARGSPSESASSAQLAPSLKA	2213
QY	2218	TCTTHGKADVDMVDAINF---WGDVTRIESSGRVVVLDLDPWEEBSDEPSPSE	2273
Db	2214	TCTINHDSPPAELEIANLMLROEMGNGITRVESEKVVILDSFPLVAEEDEREISVPAB	2273
QY	2274	YMLPKRPPALPAMARPDPNPLVESWKRPPYOCATYAGCALPPPKTTPPPRRBRTV	2333
Db	2274	ILKSRRTFOALPINARPDYNPPLLETWKKPYTEPPVNHGCLPPLPGSPPIPPRRKRTV	2333
QY	2334	GLSEDSIGALQOLAIKSFQPPPSGDSGLSTGAGADSGSQTPPELDALSETGISISMP	2393
Db	2334	VLTESTSLTALAEIAKSAFSGSSSTSGITQDNMTTTSBPAPGSGSPD---SDASEYSMP	2389
QY	2394	PLAGEGLDPDLBEQVEPQPPQGGVAABGSGSGNSCTSEB--DDSVYCCSMYSMTGA	2451
Db	2390	PLGEPEPDPPL-----SDGWSVTSVSEBAGTEDEVVCCSMYSMTVGA	2429
QY	2452	LTPGCPPEEELKPIINPLNSLRYHNKYCTTTSKASIRAKKVTFFDMQVDSYYDVLK	2511
Db	2490	EYVAAASAKYKANLISVEEACSLTPHPSAKSKFGYAKDVRCHAKAVNHINSWKDLLED	2549
QY	2572	SENPPIPTTMANKVEVCVDPRTGKKAKARLYYPLDGRVGEKALYDIQKLEQAAWGA	2631
Db	2550	SVTPIQTTIMAKNEVFCVQPEKGRKPAIIVFPLDGRVGEKALYDVVSKLEPPAWGS	2609
QY	2632	SYGFVSPAPQVRVEFLKAAVEKKDPMWGSYDTRCDSTVTRDRDIRTEBSIYRACSLPEEA	2691

Db 2610 SYGFQSYSGQVREFLVQAMKSKRTPMGFSYDTRCFDSITVTSEDIKTEEAIVCCDLDPQA 2669
QY 2692 HTAHSLTERLYVGGPMNSNGQTCGYRCRASGVLTSMKGTITCYKALACRAAGI 2751
Db 2670 RVANSLTERLYVGGPLTNSRGENCGRASGVLTSCGNTLTCTYIKARAAAGI 2729
QY 2752 APTMLVCGDDLVVISESGCTEEDERNLAFTEAMTRYSAPEGPDPREYDELITSCSSN 2811
Db 2730 DCTMLVCGDDLVVIESAGVEDASLAFTEAMTRYSAPEGPDPREYDELITSCSSN 2789
QY 2812 VSVAGPGQRRRYTTRDPTPIADAMETVHSPVNSWLGNIIOYAPTIAWMTMTF 2871
Db 2790 VSAHADGCKRYYVYTRDPTPIADAMETASHTPVNSWLGNIIMFAPTIAMRMIMTMF 2849
QY 2872 FSIIAAGTLPQNLNFEWGANYSVPLDPAIIRLGLDAFSLHTTTPHETRYASAL 2931
Db 2850 FSVLIARQLEQALDCEIYGACYSIEPLDPTIQLRLGLSAPLSHSYSPGEINRVACL 2909
QY 2932 RKLGAPELRAMKSRARAVASLISRGRAAVCGRYLFNMAVTKLKLPLPEARLLDLS 2991
Db 2910 RKLGPPLRAMRHRARSRRARLLSRGRRALICGKLFNMAVTKLKLPLPIAAGRLDLSG 2969
QY 2992 WFTVGAGGDIYHVSRRARPLLPLFGLLLFVGVGLFLPAR 3033
Db 2970 WFTAGYSGGDIYHVSRRARPRWFECLLLAAGVGIYLLPNR 3011

RESULT 9

S18030
genome polyprotein - hepatitis C virus (isolate JKI)
N:Contains: capsid protein C; envelope protein M; hepatitisin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
A:Variety: isolate JKI
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: S18030; S33570; A48332; S18029
R:Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A:Description: A whole genome of hepatitis C virus cDNA was isolated from a single patie
A:Reference number: S18028
A:Accession: S18030
A:Molecule type: genomic RNA
A:Residues: 1-3010 <HON>
A:Cross-references: UNIPROT:068949; EMBL:X61596; NID:G59478; PIDN:CAA43793.1; PID:G59479
R:Experimental source: isolate JKI from an individual
R:Honda, M.; Kaneko, S.; Unoura, M.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A:Title: Sequence analysis of putative structural regions of hepatitis C virus isolated
A:Reference number: A48332; MUID:93119270; PMID:8380322
A:Accession: S33570
A:Molecule type: genomic RNA
A:Residues: 1-547,'T',549-621,'V',623-624,'S',626-652,'D',655-761,'T',763-782 <HOM>
A:Cross-references: EMBL:X61591
A>Note: this sequence is inconsistent with the nucleotide translation
A>Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue 320
as Trp, and TTC for residue 771 as Ser
A>Note: sequence extracted from NCBI backbone (NCBIN:121747, NCBI:121748)
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; glycoprotein; hydrolyase; nucleotide binding; P-loop; polyprotein; serin
F:2-115/Product: capsid protein C #status predicted <CPC>
F:115-191/Product: envelope protein M #status predicted
F:191-389/Product: major envelope protein E #status predicted <ME>
F:389-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitisin #status predicted <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <NS4>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <NS4>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate (As

Query Match 73.5%; Score 11843.5; DB 1; Length 3010;
Best Local Similarity 70.9%; Pred. No. 0;
Matches 2157; Conservative 346; Mismatches 496; Indels 45; Gaps 8;

QY 1 MSTNKPORKTKRNTNRPODVKPPGGQIVGVYLLPRRGRGLGRATKTSERSQPRG 60
Db 1 MSTNKPORKTKRNTNRPODVKPPGGQIVGVYLLPRRGRGLGRATKTSERSQPRG 60
QY 61 RROPIDKRRSTGSKMGKPGYPMPLYGNEGIGMGWMLSPGSRPSKPNDRSRVNG 120
Db 61 RROPIDKRRSTGSKMGKPGYPMPLYGNEGIGMGWMLSPGSRPSKPNDRSRVNG 120
QY 121 KYIDLTGCFADLMGYIVVGAAPLGVARALAHGVRVLEDGVNFATGMLPGCSFESIFLLA 180
Db 121 KYIDLTGCFADLMGYIVVGAAPLGVARALAHGVRVLEDGVNFATGMLPGCSFESIFLLA 180
QY 181 LLSCTTPVSAAEVKNISITGYVMTNDCTNDSITWQLQAHLHPGVCPCEKYNASQCI 240
Db 181 LLSCTTPVSAAEVKNISITGYVMTNDCTNDSITWQLQAHLHPGVCPCEKYNASQCI 240
QY 241 PVSPPVAVORRGALTOGLRTHIDMVYMSATLCSALYVGDLCGGVTLAQMFIVSQHHWF 300
Db 241 ALTPTLAARNSSIPPTTIRRHVDLLVGAALCSAMVVDLCGSVFLVSQLETFSSRRYET 300
QY 301 VODCNCSITPGTITGHRAMDMNMMNSPTATMILAYARVBEVIDIISGAHMGMPGLA 360
Db 301 VODCNCSITPGTITGHRAMDMNMMNSPTATMILAYARVBEVIDIISGAHMGMPGLA 360
QY 361 YFSMOGAAKVVVILLLAGVDARTHTVGSAAQTGRLTSLFDMGPRQKIQLVNTNSG 420
Db 361 YFSMOGAAKVVVILLLAGVDARTHTVGSAAQTGRLTSLFDMGPRQKIQLVNTNSG 420
QY 421 HINRTALNCNDSLRHTPLASLFYTHSPFSSGCEPMSCRSIEARVCGMALQYEDNTN 480
Db 421 HINRTALNCNDSLRHTPLASLFYTHSPFSSGCEPMSCRSIEARVCGMALQYEDNTN 480
QY 481 PEDRPPCMYHPROCGVSAKTCVGPYCTPSPVVVGTDRGAPTYTGENTDVEL 540
Db 481 PEDRPPCMYHPROCGVSAKTCVGPYCTPSPVVVGTDRGAPTYTGENTDVEL 540
QY 541 LNSTRPPLGSWFGCTMNSSGYTKTCGAPPCRTADFNASTDLCLPTDCFRKHPDITYLK 600
Db 541 LNSTRPPLGSWFGCTMNSSGYTKTCGAPPCRTADFNASTDLCLPTDCFRKHPDITYLK 600
QY 601 CGSGMWLTPLRCLIDYPTLHMYCTVNTYTKIRYVGVGVEHRLTAACNFRGRDCNIED 660
Db 601 CGSGMWLTPLRCLIDYPTLHMYCTVNTYTKIRYVGVGVEHRLTAACNFRGRDCNIED 660
QY 661 RDRSOLSPLSHTTEWATLPCSYSDLPALSTGLHLHONIVDVOFMYGLSPALTYIVRW 720
Db 661 RDRSOLSPLSHTTEWATLPCSYSDLPALSTGLHLHONIVDVOFMYGLSPALTYIVRW 720
QY 721 EMWILLFLLADARVCACTMMLILGQEAALAEKLVILHAASAAACNGFLYFVIFVAVW 780
Db 721 EMWILLFLLADARVCACTMMLILGQEAALAEKLVILHAASAAACNGFLYFVIFVAVW 780
QY 781 YIKGRVPLATYSLTGLWSFLLILAPQAVAYDASVHGQIGAAALVMITLFTLPGYK 840
Db 781 YIKGRVPLATYSLTGLWSFLLILAPQAVAYDASVHGQIGAAALVMITLFTLPGYK 840
QY 841 TLLSRFLMWLYLTLGAMQEWAPRQVGGDGIIWVAIFPGVVDITKMLAVL 900
Db 841 TLLSRFLMWLYLTLGAMQEWAPRQVGGDGIIWVAIFPGVVDITKMLAVL 900
QY 901 GPAYLLKGLATRVYFYVAHALBMCMAHLAGGRVQVQMLALGRTGTYYIDHLLTPM 960
Db 901 GPAYLLKGLATRVYFYVAHALBMCMAHLAGGRVQVQMLALGRTGTYYIDHLLTPM 960
QY 961 SDMAASGIRDLAVAVPEPIIISPMKKYIWKAEALTAACDILHGIPLYSKRLGREVLGPD 1020
Db 961 SDMAASGIRDLAVAVPEPIIISPMKKYIWKAEALTAACDILHGIPLYSKRLGREVLGPD 1020
QY 1020 RDMAHSGIRDLAVAVPEPIIISPMKKYIWKAEALTAACDILHGIPLYSKRLGREVLGPD 1080
Db 1020 RDMAHSGIRDLAVAVPEPIIISPMKKYIWKAEALTAACDILHGIPLYSKRLGREVLGPD 1080

QY 1021 GYTSKMSLIAPITAYAOQTRGLLGTIVSMTRGRDTEQAGEIQVLSTVTSQSLFSTISG 1080
Db 1017 GFPEQCGMRLLAPITAYASQQTRGLFGCIVTSLTGRDKNQVEGAQVSTATQSLFATCVNG 1076
QY 1081 VLVTVVHAGNKTLAGSRGVPYTMYSABEDVGMPSRPGCTKLEPCTCTGAVDLVYTMN 1140
Db 1077 VCVTVVHAGSKTLTAGKGPINQMVTYNVDDLVGMQAPSGAASLTCTYGGSDLVYVTH 1136
QY 1141 ADVIPARRRGDGRGALLSPRLSTLKSGSSGPLYLCRGAHVGFRAAVCSRGVAKSIDPI 1200
Db 1137 ADVIPARRRGDSRGLSPRPVSYLKSSSGPLLCRSGAHVGFIRAVCTCRGAKAVDPI 1196
QY 1201 PVETLDVTRSPLTFSDNSTPRAVPQYQVGYLHAPSGSKSTVPVAYAAQYKVLVNP 1260
Db 1197 PVESMETTRMSPVFTDSSSPPAVQTFQVNHLPAGSGSKSTKVPAPVAAQYKVLVNP 1256
QY 1261 SVAAITLGFAYLKAHGININRTGVRTVTGAPIYSTYTKGLADGGCAGAYDIIID 1320
Db 1257 SVAAITLGFAYMKGAVNDPNISTGVRTITTGAPITYSTYTKGLADGGCAGAYDIIID 1316
QY 1321 ECHAVDSTTLIGITVLDQAEIAGVRLTVLATATPPGSVTTPHPNIEVALGQEGEIPY 1380
Db 1317 ECHSTSTSLIGITVLDQAEIAGVRLTVLAATPPGSVTPHHPNIEVALPMTGSEIPY 1376
QY 1381 GRAIPLSYIKGSHLIFCHSKKKCDLAAALRGKLSNVAAYRGDLVSVIPTQGDVVVA 1440
Db 1377 GKAIPLETIKGRHLLIFCHSKKKCDLAAALRGKLSNVAAYRGDLVSVIPTSGDVVVA 1436
QY 1441 TDALMTGYTDPDSVIDCNVAVTQVVDPSLDPTFTITTOVPODVAISRGRRTGRGL 1500
Db 1437 TDALMTGYTDPDSVIDCNVTQVVDPSLDPTFTITTOVPODVAISRGRRTGRGL 1496
QY 1501 GIYRVYSTGBRASGMPDSVVLCECYDAGAAWELTPSETTVRLRAYFNTPGJLVCCDHE 1560
Db 1497 GIYRVFTGSRPGSGMPDSVVLCECYDAGAAWELTPRAYVSRRLAYLNTPGJLVCCVHLE 1556
QY 1561 FMAVAVFTGLTHIDAHFLSOTKSGENPALTAYQATYCARAKAPPSQWDMVKCLTRLP 1620
Db 1557 FMSVSVFTGLTHIDAHFLSOTKSGENPALTAYQATYCARAKAPPSQWDMVKCLTRLP 1616
QY 1621 TLVGPPLRYRLGSVTNEVTLTHPVTKYIATCMQADLEWMTSTWVLAVGLVLAAYVCLA 1680
Db 1617 TLVGPPLRYRLGAVQNEVTLTHPVTKYIATCMQADLEWMTSTWVLAVGLVLAAYVCLT 1676
QY 1681 TGCVCIIIGRLHINQRAVAVADKEVLYEAFDEMECASRAALIEBEGRIEMLSKIQGL 1740
Db 1677 TGSVVIIVGRILILSGRAPIIPDREVLYOEPDEMECASHLPIEQMGQLAQFOKXGL 1736
QY 1741 QOASKQAOIQIPVQASWPKVEQFMAKMMNFTISGIOYLAGLSTLPQNPAVSMMAFSA 1800
Db 1737 QOASKQAEAAAPVYESKQALBEAFMAKMMNFTISGIOYLAGLSTLPQNPAVSLMAFTAS 1796
QY 1801 LTPPLSTSTTLINILGMLASQIAPAPGATGVVSGVLAAGVSGIGLGVLVNDILAGV 1860
Db 1797 LTPPLTQHTLNLNIIAGMVAQIAPPSAASAFVAGIGAGAAVSGIGLGVLVNDILAGV 1856
QY 1861 AGISGALVAEKINGSEKPSMEDVYVNLPGILSPGALVGVICAAILRRHVPEGEAVQM 1920
Db 1857 AGVAGALVAEKMGSGEMPTEDLVNLPAITLSPGALVGVCAAILRRHVPEGEAVQM 1916
QY 1921 NRIIAPASRGMYAPHTYTESDASQRYVQLLSLTITSLRLHLMWITEDCIPGSGW 1980
Db 1917 NRIIAPASRGMYAPHTYTESDASQRYVQLLSLTITSLRLHLMWITEDCIPGSGW 1976
QY 1981 LRDVMDVAVCTIILNDFKNMLTSKLPKMBGLPFVSCQGYGVAGTGMTRCPGANIS 2040
Db 1977 LRDVMDVAVCTIILNDFKNMLTSKLPKMBGLPFVSCQGYGVAGTGMTRCPGANIS 2036
QY 2041 GNVRLGSMRITGKTCMNIIQGTFFPINCYTEGQCVKPAENFVAIMRWVAASEYAEVTH 2100
Db 2037 GHVXNGSMRIVGPKTCMNIIQGTFFPINCYTEGQCVKPAENFVAIMRWVAASEYAEVTH 2096
QY 2101 GSHYIITGLTTDLNKLKVPCCQLPSPBPFMSWVDGVQIHRPAPTRPKFPFDDVSVFCGLNSPV 2160

Db 2097 GFHYVYMTTDDNVKCPQVAPAPREFTEVDVRLHRYVPAKCPRLRDEVTQVGLNPFV 2156
QY 2161 GSQLPCEDEPDVDVMSMLTDPSSH TAETARRLARGSPSEASSASQASPSLRATCT 2220
Db 2157 GSQLPCEDEPDVDVMSMLTDPSSH TAETARRLARGSPSEASSASQASPSLRATCT 2216
QY 2221 THGRAVDVMDANLF---MGDVTRIESGSKVYVLDLSDPMVEERDLEPSIPSEYML 2276
Db 2217 TRHSDPADLLEBANLWMEKGNITRVESNKNVYILDSFPLRABEDREVSVAELIR 2276
QY 2277 PKRFPPLPAMARPDVNPVLVESKRPDYOPATYVAGCALPPEPKTPPPRRRTVLS 2336
Db 2277 KSRKFPPLPAMARPDVNPVLVESKRPDYOPATYVAGCALPPEPKTPPPRRRTVLT 2336
QY 2337 EDSIGDALQOALISFGQPPSGDSGTAGAADSGQT--PPDELA-----LSBTGIS 2390
Db 2337 ESTVSALAEIATYKTFG-----SSGSANDSGTATAPPDQPSDDGSGSDDEYS 2386
QY 2391 SMPPLEGBLGDPLEPEQVEBPQPPQGVAAPGSDSGSWSTCSEB-DDSVVCSGMSYMT 2449
Db 2387 SMPPLEGBLGDPLEPEQVEBPQPPQGVAAPGSDSGSWSTCSEB-DDSVVCSGMSYMT 2426
QY 2450 GALTTPCSPSEBCKLPINPLNSLRLRYHNKYCTTTKASLSAKKYTFPRMQLVLSYVS 2509
Db 2427 GALTTPCAABESKLPINPLNSLRLRYHNKYCTTTKASLSAKKYTFPRMQLVLSYVS 2486
QY 2510 LKDIKLAASKYARLLTWEEACOLTPPHSASKYGFGEKAYRSLSGRAVNHAKSYWKDL 2569
Db 2487 LKEMKAKASTYAKLLSYEBCKLTPPHSANSKYGAKADVRANSSKAVNHHSYWKDL 2546
QY 2570 EDESETPIETIMAKNEVCVDPTRKGGKKAABLYVVDLGVVCEKMALYDITQKLPQVM 2629
Db 2547 EDESETPIETIMAKNEVCVDPTRKGGKKAABLYVVDLGVVCEKMALYDITQKLPQVM 2606
QY 2630 GASVGFQYSPARVBEFLKAAEKDPMGSEYDTRCPDSTYTEEDIRFEESTIYACSLPE 2689
Db 2607 GSSVGFQYSPARVBEFLKAAEKDPMGSEYDTRCPDSTYTEEDIRFEESTIYACSLPE 2666
QY 2690 EAHYTHSLTBLRYVCGPMFNSKQOTCYRRCRASGVLTSMGNTITCYVLAALAAKRAAG 2749
Db 2667 EAHYTHSLTBLRYVCGPMFNSKQOTCYRRCRASGVLTSMGNTITCYVLAALAAKRAAG 2726
QY 2750 IIAPTMLVCGDVLVVISOSGTEBEDRLNRAFTAMRTYSAPGDPPEPEYDELTISCS 2809
Db 2727 IODCTMLVCGDVLVVISOSGTEBEDRLNRAFTAMRTYSAPGDPPEPEYDELTISCS 2786
QY 2810 SNVSVALDPQGRRYVLTTRDPTPIAARAAMETVHSPNSWLNINIOYAPITWARMVLT 2869
Db 2787 SNVSVALDPQGRRYVLTTRDPTPIAARAAMETVHSPNSWLNINIOYAPITWARMVLT 2846
QY 2870 HFFSILMAQDTLDQNLNPEMGAIVSVPLDPAIIEHLGLDAFSLHTYPPHELTRVAS 2929
Db 2847 HFFSILMAQDTLDQNLNPEMGAIVSVPLDPAIIEHLGLDAFSLHTYPPHELTRVAS 2906
QY 2930 ALRKLGADPLRAWMSRADAVALSLISRGRAAVCGRYVFNVAVYTKLPLPEARLIDL 2989
Db 2907 CLRKLGADPLRAWMSRADAVALSLISRGRAAVCGRYVFNVAVYTKLPLPEARLIDL 2966
QY 2990 SSMFTVGAAGGDIVHSVRARPLLFGLLLPFGVGLFLPAR 3033
Db 2967 SSMFTVGAAGGDIVHSVRARPLLFGLLLPFGVGLFLPAR 3010

RESULT 10
GNMWH
genome polypeptide - hepatitis C virus (strain H)
N:contains: capsid protein C; envelope protein M; hepatitisin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: A36814; A41546

R:inchaupae, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
submitted to GenBank, July 1992
A:Description: Genomic structure of the human prototype strain H of hepatitis C virus: C
A:Reference number: A36814
A:Accession: A36814
A:Molecule type: genomic RNA
A:Residues: 1-3011 <INC>
A:Cross-references: UNIPROT:p27958; GB:M67463; NID:G329737; PID:AAA5534.1; PID:G329738
R:inchaupae, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
Proc. Natl. Acad. Sci. U.S.A. 88, 10292-10296, 1991
A:Title: Genomic structure of the human prototype strain H of hepatitis C virus: compari
A:Reference number: A1546; MUID:92052256; PMID:1658800
A:Contents: annotation
A:Note: neither amino acid nor nucleotide sequence is given
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura
F:1-115/Product: capsid protein C #status predicted <CPC>
F:115-191/Product: envelope protein M #status predicted <EBM>
F:191-389/Product: major envelope protein E #status predicted <NS1>
F:389-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus nucleotide-binding motif A (P-loop)
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <NS4a>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <NS4b>
F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F:196, 209, 234, 305, 325, 417, 423, 430, 448, 476, 532, 540, 556, 576, 623, 645, 1213, 1255, 2041, 2240, 23

Query Match 73.2%; Score 11796; DB 1; Length 3011;
Best Local Similarity 70.7%; Pred. No. 0;
Matches 2149; Conservative 364; Mismatches 491; Indels 36; Gaps 8;

QY 1 MSTNPKPQRTKRNTRRPODVKPGGQIVGVYLLPRGRRLCVARTKTSERSQPRG 60
DB 1 MSTNPKPQRTKRNTRRPODVKPGGQIVGVYLLPRGRRLCVARTKTSERSQPRG 60
QY 61 RROPFPKRRSTGSKGKPGYWPPIYGNBGLGMAWLLSPGSRSSKGNPNRHRNRVNG 120
DB 61 RROPFPKRRSTGSKGKPGYWPPIYGNBGLGMAWLLSPGSRSSKGNPNRHRNRVNG 120
QY 121 KVIDTLTGCPADLMGYIPVNGAPLGGVABALAHGVRVLEQVNFATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTGCPADLMGYIPVNGAPLGGVABALAHGVRVLEQVNFATGNLPGCSFSIFLLA 180
QY 181 LLSCTTTPVSAAEVKNISTGVWVNTDCTNDSITWQLQAAVLHPGCVPCCEKVGNSQCM 240
DB 181 LLSCTTTPVSAAEVKNISTGVWVNTDCTNDSITWQLQAAVLHPGCVPCCEKVGNSQCM 240
QY 241 PVSPNVAVORPGL-TQGLRTHIDMVMSATLCSALYYGDLCCGVMLAQMFTVSPQHH 299
DB 241 PVSPNVAVORPGL-TQGLRTHIDMVMSATLCSALYYGDLCCGVMLAQMFTVSPQHH 299
QY 300 FVQDNCGIYPTGTHGHAMDMNMWSPATMILAAVMPVEVIDIISAGHMCVMFCL 359
DB 300 FVQDNCGIYPTGTHGHAMDMNMWSPATMILAAVMPVEVIDIISAGHMCVMFCL 359
QY 360 AYFSQAGAAKAVVILLAAVDARTHTVVGSAQTGRLTSLFDMGPQKIQLVNTG 419
DB 360 AYFSQAGAAKAVVILLAAVDARTHTVVGSAQTGRLTSLFDMGPQKIQLVNTG 419
QY 420 WHINRTALNCDSLHTGFIASLFYTHSRSSGCCPERMSACRSIAEPRGKMLQYEDVNT 479
DB 420 WHINRTALNCDSLHTGFIASLFYTHSRSSGCCPERMSACRSIAEPRGKMLQYEDVNT 479
QY 480 NPEDMRPYCMHYPQCGVSAKTVCGPVYCTPSPVVVGTDRDAGAPTYTGESEITDF 539
DB 480 NPEDMRPYCMHYPQCGVSAKTVCGPVYCTPSPVVVGTDRDAGAPTYTGESEITDF 539
QY 540 LINSTRPPLGSMFGCTMANSSTGYTTCGAPCPTRADFNASTDLLCPTDFCRKHPDPTYL 599
DB 540 LINSTRPPLGSMFGCTMANSSTGYTTCGAPCPTRADFNASTDLLCPTDFCRKHPDPTYL 599

QY 600 KCGSGPMLTPPCRLIDYPYRLMHPYPCVTNVTIFKIRMYGVGEHRLTAACNFRGCRNLE 659
DB 596 KCGSGPMLTPPCRLIDYPYRLMHPYPCVTNVTIFKIRMYGVGEHRLTAACNFRGCRNLE 659
QY 660 DDRSGLPPLLSTTEWALILPCSYSDLPALSTGLHLHQNIVDVQPMYGLSPALTKYIVR 719
DB 656 DDRSGLPPLLSTTEWALILPCSYSDLPALSTGLHLHQNIVDVQPMYGLSPALTKYIVR 719
QY 720 WENYVLLFLLLADARVCCIMMLILGQAEALLETULILHMAASACNGFLYFVLPFAVA 779
DB 716 WENYVLLFLLLADARVCCIMMLILGQAEALLETULILHMAASACNGFLYFVLPFAVA 779
QY 780 WYIKRVVPLATYSITGLMSFSLLLLAIPQOAVYADASVHGOIGALVMTILFTLTPGY 839
DB 776 WYIKRVVPLATYSITGLMSFSLLLLAIPQOAVYADASVHGOIGALVMTILFTLTPGY 839
QY 840 KTLASRFLMWLCYLLITLGEAVVQEWAPMQRGRDGIIMAVAFYPGVVDITKMLAV 899
DB 836 KRYISMCMWMLQYFLTRVEAQLHWVPPPLNVRGRDAVILLTCVHPALVFDITKLALAI 895
QY 900 LGPAILKGLATRVYFPRALALLMCTMAHRLAGRVQNALALLGRMTGYITDHLTP 959
DB 896 LGPAILKGLATRVYFPRALALLMCTMAHRLAGRVQNALALLGRMTGYITDHLTP 959
QY 960 MSDMASGLRDLAVAVEPIIISPMKRYIVWGAETPAACGDIHLGLPVASRLGREVLGLPA 1019
DB 956 LRDMAHNRGRDLAVAVEPIIISPMKRYIVWGAETPAACGDIHLGLPVASRLGREVLGLPA 1015
QY 1020 DGYTSKMSLAPITAYVAAQOTRGLLGIIVSWMTGRDKTEQAGEIQVLSVTQSFILGTSIS 1079
DB 1016 DGYTSKMSLAPITAYVAAQOTRGLLGIIVSWMTGRDKTEQAGEIQVLSVTQSFILGTSIS 1075
QY 1080 GYLATVYHGAENKTLIAGSRGYTQMSAEGDLYWMBPBGTKSLPECTCGAVDLYLVR 1139
DB 1076 GYLATVYHGAENKTLIAGSRGYTQMSAEGDLYWMBPBGTKSLPECTCGAVDLYLVR 1135
QY 1140 NADVPARRRDPKRGALLSPRLSTLKGSSGGPVLCPRGHVGVVRAVACSRSKSIDF 1199
DB 1136 NADVPARRRDPKRGALLSPRLSTLKGSSGGPVLCPRGHVGVVRAVACSRSKSIDF 1195
QY 1200 IPVETLIDIVTSPFESDNTSPPAVQYOVQVGLHAPTSKSTKVPVAAAGYRVLVN 1259
DB 1196 IPVETLIDIVTSPFESDNTSPPAVQYOVQVGLHAPTSKSTKVPVAAAGYRVLVN 1255
QY 1260 PSVAATLGFAYLSKAGINPNIQTGRTVTTGABITYSTYTKFLADGGCAGAYDIIC 1319
DB 1256 PSVAATLGFAYLSKAGINPNIQTGRTVTTGABITYSTYTKFLADGGCAGAYDIIC 1315
QY 1320 DECAVDSITLIGCTVLDQAEATGVRLTVLATATPPGSAVTPPHNIEBALGOGGEI 1379
DB 1316 DECAVDSITLIGCTVLDQAEATGVRLTVLATATPPGSAVTPPHNIEBALGOGGEI 1375
QY 1380 YGRAIPLSYIKGRHLIFCHSKKKCCDELAALRGKGLNSVAYYRGLDVSIPTQGDVVV 1439
DB 1376 YGRAIPLSYIKGRHLIFCHSKKKCCDELAALRGKGLNSVAYYRGLDVSIPTQGDVVV 1435
QY 1440 ATDALMTGYTGFDSVIDCNVAVTVQVDFSLDPTFTTTQVLPQDAVRSQRRGTGGR 1499
DB 1436 ATDALMTGYTGFDSVIDCNVAVTVQVDFSLDPTFTTTQVLPQDAVRSQRRGTGGR 1495
QY 1500 LGIYRYSVTSGRASGMPDSVYLCECYDAGAAWYELTPESETVRLRAYENTPGLPVCODHL 1559
DB 1496 LGIYRYSVTSGRASGMPDSVYLCECYDAGAAWYELTPESETVRLRAYENTPGLPVCODHL 1555
QY 1560 EFWMAVFTGLTHIDAHFLSQTQSGENPAVYLJAYQATVCAKAKAPSPMDVMWKLTRLK 1619
DB 1556 EFWMAVFTGLTHIDAHFLSQTQSGENPAVYLJAYQATVCAKAKAPSPMDVMWKLTRLK 1615
QY 1620 PTLVGPFTLLYRLSGVYNEVLTHTVTKYIATQCADLEWTSSTVYLAGVLAANAAYCL 1679
DB 1616 PTLVGPFTLLYRLSGVYNEVLTHTVTKYIATQCADLEWTSSTVYLAGVLAANAAYCL 1675

Db 181 LLMCLTVBASAVPYENNASGVHTVNDPCNSSIYVEADNLLIHPGCVPLEEDNYSRCWV 240
 QY 241 PVSVPVAVORPQALQOGLRTHIDMVMVMAATLCSALUYGDLGGWLLAQMFTVSPQHMF 300
 Db 241 QITPLUSPFSFAVATLRRADVYLAGGAFCALYVGDCALSLVGMFTYKRRQHT 300
 QY 301 VODCNCSIYPCGTTGHRAMDMNMMSPTATMILAYAMRPEVIIDISGAMVWFGLA 360
 Db 301 VODCNCSISGHTTHRAMDMNMMSPTATMILMQLRIPQVVIDIAGSMGVLMAA 360
 QY 361 YFSMGAMAKVVVILLAAVDARTHTVGSAAQTTRLTSLFDMGPROKIOLVNTNSW 420
 Db 361 YEASTANNAKVLIVFLFAGVDGRTHTVGGIGGLSLTSPFNCGPQROLOFVNTNSG 420
 QY 421 HINRTALNNDLSHGTFLASLFYTHSPNSGGCEBMSACRSLEARRVGMALQYEDNVTN 480
 Db 421 HINRTALNNDLSQGLFAGLMTYAHKFNSSGCEBMSACRPLAADQGGITISYA-TISG 479
 QY 481 PEDMBPYCWHYPPROCGVVSATKVCSPVYCTPSPVVVGTTRDLGAPTYMGENTDVEL 540
 Db 480 PBDDBPYCWHYPPRCGVVPADVCGPYVCTPSPVVVGTTRDRCGPTVMKSNETDILL 539
 QY 541 INSTRPPLGSMFGCTWMNSGYTKTCGAPPCRTADFNASTDLCPTDCFRHPTTYLK 600
 Db 540 INNIRPPAGNMFGCTWMNSTGFEVKNCGAPPCNIGFTGNNS--LKCPTDCFRHPTATYTR 597
 QY 601 CGSGWMLTPRCLIDVPRYLMHYPCVTNVTIFKIRMYGVGVHRLTAACNFTGDCNLED 660
 Db 598 CGSGWMLTPRCLVHYPRYLMHYPCVTNVTIFVYRMEIGGLEHRLBAACNWTYGERCDLED 657
 QY 661 RDRSOLSPLLHSTTEMALPCSYSDLPALSTGLHLHONIVDQFMYGSLPALTKYIYK 720
 Db 658 RDRALSLPLHTTTQWAILPCSFTEPRALSTGLHLHONIVDQVLYGSLSSIVEMAYKM 717
 QY 721 EYVILLFLILADARVACILMLILGQEALEKLVILHAASASCNGLFYVIFVVAAM 780
 Db 718 EYIMVFLILADARICTCLILILILGQEAATCKNIVILNAAAAAGNHGFFMGLVVCIAM 777
 QY 781 YIKGVNPLATYSLTGLMSFSLLLALPOAYAYDASVHGOLGALLVMTLFTITPGK 840
 Db 778 HVGKGLVPCATYLCIGWPELLVRLRLRPRALALDSSDGTGCVLVLTLFTITPGK 837
 QY 841 TLLSFLWMLCYLLTLGEMVOEMAPPMOVGRGDIWAVALIFPGVVFDTIKMLAVL 900
 Db 838 KKVVLVMMLOQYFIRAVEAIIHWWPPLQYKGRDAVIMTCLFPRALGFEITIKLFGIL 897
 QY 901 GPAYILLKGLATREVPYFVRAHALLRMCMAHLAGRYYVQMALLAGRWTYIYDHLTPM 960
 Db 898 GPLYLLQHSLSLTQVYFELRABALLRLCLAKHLVYGKYOALHLHGRLTGYIYDHLAPM 957
 QY 961 SPMAASGLDLAVNEPIIFSPMEKKVIVMGAEATAAGDILHGLVYSARLGREVLGPRD 1020
 Db 958 KDMAASGLRELVAATEPIVFSAMETKVITWADTAACNILLAVLPSARGEIIFLGPAD 1017
 QY 1021 GYTSKMSLILATITVAOQTRGLGTIVYSMTGRDTEOAGEIOVLSYTYOSFLGTSISG 1080
 Db 1018 DIKTSKMSLILATITVAOQTRGLGTIVLSLGRDKNAEAGEVOQLSTATQYFLGICING 1077
 QY 1081 VLMTVYHAGANKTLGASRGVPTYOMYSAAEGDLVGMPSPEPTKSLPCTCGADVLYLVRN 1140
 Db 1078 VMMTLFPHGAGSKTLGAPKGPVVQMYTNDKDLVGMPSPPGKSLTRCTGSAADLYLVTTH 1137
 QY 1141 ADVIARRRGDKRGLLSRPLSTLKSGSGGPRVLCPRGNAVFPRAAVCSRVAASIDPI 1200
 Db 1138 ADVIARRRGDTRASLLSRPISYLGSGSGGPRIMCPSGHVGVFPRAAVCTRRVAAALEV 1197
 QY 1201 PVETIADIVRSPTFSDNSTPPAVPQTYGYVYLHAPTSKSTKVVAAVAAQGYKVLVNP 1260
 Db 1198 PVENIETTRMSRVTFDNSTPPAVPHEFOVGHILHAPTSKSTKVVAAVAAQGYKVLVNP 1257
 QY 1261 SVAAATLFGGAYLSKAGHINPNIRTVGRTVTTGAPITTSYTGFLADGCGAGAYDIICD 1320
 Db 1258 SVAAATFGGAYMSRAYVDPNIRITGVRTVTTGAGITTSYTGKFPADGCGAGAYDIICD 1317

QY 1321 ECHAVDSPTIIGITVLDOAETAGVRLTVLATAPRPGSVTTPHPNIEBALGOGEIIPFY 1380
 Db 1318 ECHSODATTIIGITVLDOAETAGVRLTVLATAPRPGSVTTPHPNIEBALPSEIIPFY 1377
 QY 1381 GRAIPLSYIKGRHLIFCHSKKKCDELAALRGMLNSVAAYYRGADVSIPTOGDVVVA 1440
 Db 1378 GRAIPLVLIKGRHLIFANQKKKAKETAKONMKPKRAKAAVAYRGADVSIPTOGDVVVA 1437
 QY 1441 TDALMTGTPGPDSDYIDCNVAVTQVVDSPDPTFTTQIYIPQDVAVSSORGRGTRGL 1500
 Db 1438 TDALMTGTPGPDSDYIDCNVAVTQVVDSPDPTFTTQIYIPQDVAVSSORGRGTRGL 1497
 QY 1501 GIYRVYSGERASGMFDSVULCECYDAGAAVYELPSESTYRLRAYFNTPGLPVCOHDLE 1560
 Db 1498 GIYRVYSGERASGMFDSVULCECYDAGAAVYELPSESTYRLRAYFNTPGLPVCOHDLE 1557
 QY 1561 FWEAVFTGLTHIDAHFLSQTKOSGENFAYLTAQATVCARAKAPPSPSDVMKCLTRLPK 1620
 Db 1558 FWEAVFTGLTHIDAHFLSQTKOSGENFAYLTAQATVCARAKAPPSPSDVMKCLTRLPK 1617
 QY 1621 TLVGPPLLYRLGSTNEVTLTHPYTKIATCMQADLEVMSTWVLACGVLAANAAYCLA 1680
 Db 1618 TLVGPPLLYRLGAVQNEITLTHPYTKIYMACMSADLEVITSTWVLACGVLAANAAYCLA 1677
 QY 1681 TGVCIIIGRLHINORAVVAPDKEVLYEAFDMEBECASPAALIIEGORIAEMLSKIOGLL 1740
 Db 1678 VGSVAIVRITILSGRPAITPDBREVLYQDFDEMEBECASLPYVDEKRAIAGCFKEVYGLIT 1737
 QY 1741 OQASKOADIQPTVOASMPKVEQFWMKAMNFIISGIOYLAGLSTLPGNPAVAMMAFSA 1800
 Db 1738 GTAGOKAETLPAATISMMSKAEQFWMKAMNFIISGIOYLAGLSTLPGNPAVATLMSFTA 1797
 QY 1801 LTPSLSTTILNLTGMSLASQIAPRAGAGFPVYSGLVGAAGVSGIGKYLVDILAGY 1860
 Db 1798 VTSPLTQOTLLENLIGMSVASQIAPRAGAGFPVYSGLVGAAGVSGIGKYLVDILAGY 1857
 QY 1861 AGISALVAFKIMSGEKPMSDVNVLPGILSPGALVYGVCAAILRRHVGEAGAVOMM 1920
 Db 1858 TGVAGALVAFKIMSGEKPMSDVNVLPGILSPGALVYGVCAAILRRHVGEAGAVOMM 1917
 QY 1921 NRLIAFASRGNHVAPTHVYTESDASORVTOLLSLTLSRLRHNMTTEDCPJPCGSM 1980
 Db 1918 NRLIAFASRGNHVAPTHVYTESDASORVTOLLSLTLSRLRHNMTTEDCPJPCGSM 1977
 QY 1981 LRDVDMVCTIITDFKMLISKLPPKMGLEFPVSCQKGYKGVMACTGIMTRPCGANIS 2040
 Db 1978 LRAIWDVCTIITDFKMLISKLPPKMGLEFPVSCQKGYKGVMACTGIMTRPCGANIS 2037
 QY 2041 GNVRLGSMRITGPCKOMIWOCTPINCYTGOCVPRKAPKVAIYRVAASEVAEYOH 2100
 Db 2038 GNVRLGSMRITGPCKOMIWOCTPINCYTGOCVPRKAPKVAIYRVAASEVAEYOH 2097
 QY 2101 GSYHYITGLTNDNLKVPOLPSPPEFFSVWVDVOIHRPAPTRKPPFRDVSFCVGLNSFV 2160
 Db 2098 GSYHYITGLTNDNLKVPOLPSPPEFFSVWVDVOIHRPAPTRKPPFRDVSFCVGLNSFV 2157
 QY 2161 GSOLPCEDEPPTDVLMSLTPSPHTAETAARLARLARGSPSEBASASAOQLNAPSILACT 2220
 Db 2158 GSOLPCEDEPPTDVLMSLTPSPHTAETAARLARLARGSPSEBASASAOQLNAPSILACT 2217
 QY 2221 THGKAYVDVMDANF-----NGGDVTRIESGSKVVVLDLPMVBERSDLEPSIPEYML 2276
 Db 2218 THGKAYVDVMDANF-----NGGDVTRIESGSKVVVLDLPMVBERSDLEPSIPEYML 2277
 QY 2277 PKKRPALPAMARBDVNPPLVESKWRBDYOATVAGCALPBPRTKTPPPRR-RTYGL 2335
 Db 2278 KGAAPPALPAMARBDVNPPLVESKWRBDYOATVAGCALPBPRTKTPPPRR-RTYGL 2337
 QY 2336 SEDSIGDALQOLAIKSFQPPPS---GDSGLSTAGAADSSQTPPDLALSEFCSISSM 2392
 Db 2338 SEDSIGDALQOLAIKSFQPPPS---GDSGLSTAGAADSSQTPPDLALSEFCSISSM 2394

QY 2393 PLBSEIAGDPPLEPPEQVPEPQPPPGCGVAAAPGSDGSSNATCEEDSDVYCCSMYSMTGAL 2452

Db 2395 PPLBSEPDPPD-----SSGSMSTVSGE-DNVVCCSMYSMTGAL 2433

QY 2453 ITPCGPEEEKLPINPLSNLSLRHYNKVCTTTKSASLAKKVTDPDMOVLSYDSDVYKD 2512

Db 2454 ITPCGAEEKLPINPLSNLSLRHNLVYSTSRASGLQKKVTPRLDVLDDHREVDE 2493

QY 2513 I KLAASKVTARLLTMEAACOLTPPHSASKYFGAKEVRSLSGRAVNHKISVMWDLDEDS 2572

Db 2494 MKRLASVKYARLLPLEAACGLTPPHSASKYGAKVEVRSLDKRLAKHIEGWQDLDDSD 2553

QY 2573 ETPIPTITMAKNEVCVPTPKGKKAARLIYPLDGVAVCEKMAIYDTQKLPQAWGAS 2632

Db 2554 DTPLEPTITMAKNEVAVPESKGGKPARLIYPLDGVAVCEKRALYDVAOKLPALMSPS 2613

QY 2653 YGFOVSPARVVEFLKAAAEKDPMGFSYDTRCPSTVTERDIRTESIYRACSLPEBAH 2692

Db 2614 YGFOVSPARVDLFLKAAKSKKIPMAFSYDTRCPDSTITTEHDIMTESIYQSCDQEPAR 2673

QY 2693 TAIHSLTERLYVGGPMFNSKGGTCGRRCPRASGVLTSMGNTITCYVALAACRAAGITA 2752

Db 2674 VAIRSLTQRLVCGGPMNNSKGGCCGRRCPRASGVLTSMGNTMTCYIALASCPAAKLRD 2733

QY 2753 PTMLVGSDLVVISGCGTEEDERNLRAFTEAMTRYSAPEGDPPEYDELEITSCSSNV 2812

Db 2734 CTLLVCGDVLVAICSGQTHEDEASLRAFTEAMTRYSAPEGDPVPAYDLELVTSCSSNV 2793

QY 2813 SVALGPPGRRRYLLTRDPTTPPARAAETVAKHSPNLSNLGNIQVAPITAMRMVLMTHF 2872

Db 2794 SVAARDASNRIYYLLTRDPQVPLAKAAETAKHSPNLSNLGINIYAPITAMRIYLMTHF 2853

QY 2873 SILMAQDITLDNLNEMPGAVVSVSPLDPAIIEELHGLDAFSLHTTYPHLETIVASALR 2932

Db 2854 SYLQSQEQLEKTLAEFMGYSVTPDLPAIIQLHLGSAFSLHSTSPSEINRYASCLR 2913

QY 2933 KLGAAPPLAMKSRARAVASLSIRGGAACVGRYLFNNAVTKLKLTPLEAPRLDLSW 2992

Db 2914 KLGVPPPLAMRRARAVAKKIIAGCGRAALIGIYLFNNAVTKRLTPLEADADRDLSSW 2973

QY 2993 FTVGAGGSDIYHSVRAPRLILFGLLLFVGVGLFLLPAR 3033

Db 2974 FTVGAGGSDIYHSMRSRAPRLILCLLLLSVGVGIFLLPAR 3014

RESULT 12

Q00881

genome polypeptide - hepatitis C virus (strain J6) (fragments)

N:Contains: NS5 protein

C:Species: hepatitis C virus

C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004

C:Accession: J00881

R:Okamoto, H.

submitted to JIPID, January 1991

A:Description: The 5'-terminal and 3'-terminal sequences of the genomic RNA of hepatitis C virus

A:Reference number: J00879

A:Accession: J00881

A:Molecule type: genomic RNA

A:Residues: 1-874 (OKA)

A:Cross-references: UNIPROT:O7LZV5

A:Experimental source: strain J6

C:Superfamily: hepatitis C virus genome polypeptide

C:Keywords: polypeptide

F:150-874/Product: NS5 protein (fragment) #status predicted <NS5>

[illegible]

QY	61	RROPIPKDRBSTGSKMGKPGYPMWLYXNBEIGIAGMGMLISPRGSGRPSMGPNDRHRSRVG	120
Db	61	RROPIPKDRBSTGSKMGKPGYPMWLYXNBEIGIAGMGMLISPRGSGRPSMGPNDRHRSRVG	120
QY	121	KVIDITLTGCFADLMGXITPVVGAPLGGVBARALAHGVRVLEDDGVNATGNLPGCSHSIFLLA	180
Db	121	KVIDITLTGCFADLMGXITPVVGAPLGGVBARALAHGVRVLEDDGVNATGNLPGCSHSIFLLA	180
QY	181	LLSCITTPVSAAEYKNI STGYMTNDCTNDSITWLOQA VLVHBCVCEKVGNAOCMI	240
Db	181	LLSCITTPVSAAEYKNI STGYMTNDCTNDSITWLOQA VLVHBCVCEKVGNTSRMI	240
QY	241	PVSPRVAVQVRGALTGGIRTHIDMVVMSATLCSALHYGDLCCGVMMAOMITVSPOHMF	300
Db	241	PVSPRVAVQVRGALTGGIRTHIDMVVMSATLCSALHYGDLCCGVMMAOMITVSPOHMF	300
QY	301	VODNCSTYPTGTGHRAMDMMMMSPTATMI LAYAMRVEVITIDISGHHGVMFGLA	360
Db	301	VODNCSTYPTGTGHRAMDMMMMSPTATMI LAYAMRVEVITIDISGHHGVMFGLA	360
QY	361	YFSMOGAMAKVVI LLLAAGVDARTHTVGSAAQTGRLTSLFDMGPROKIOLVNTNGSW	420
Db	361	YFSMOGAMAKVVI LLLAAGVDARTHTVGSAAQTGRLTSLFDMGPROKIOLVNTNGSW	420
QY	421	HINRTALNCDSLHTGFIASLFTYHSRNSGCCPRMSACRSIEAFRGMGLQYEDVNTN	480
Db	421	HINRTALNCDSLHTGFIASLFTYHSRNSGCCPRMSACRSIEAFRGMGLQYEDVNTN	480
QY	481	PEDRBPYOMWHPROCGVVSATKTVGCPYCTBPVVVGTDRIGAPTYTGAEHTDVL	540
Db	481	PEDRBPYOMWHPROCGVVSATKTVGCPYCTBPVVVGTDRIGAPTYTGAEHTDVL	540
QY	541	LNSTRPPLGSMFGCTMMNSGYTCTCGAPCPTRADFNASTDLCTPTDCFRKHPDTTYLK	600
Db	541	LNSTRPPLGSMFGCTMMNSGYTCTCGAPCPTRADFNASTDLCTPTDCFRKHPDTTYLK	600
QY	601	CGSGPWLTTPRCLIDYPRYLMWHY PCTVNTYTFKIMYVGVGEHRLTAACNFTGRDNCLED	660
Db	601	CGSGPWLTTPRCLIDYPRYLMWHY PCTVNTYTFKIMYVGVGEHRLTAACNFTGRDNCLED	660
QY	661	RDRQSLSPLHSTTBMAILPBCSYDLPALSTGTLHQNIVDQPMYGLSPALTKYIVRW	720
Db	661	RDRQSLSPLHSTTBMAILPBCSYDLPALSTGTLHQNIVDQPMYGLSPALTKYIVRW	720
QY	721	EMVILLFTLLADARVCAIMMLILLOQEALEKLVILHAASASCNGLFYFVIAW	780
Db	721	EMVILLFTLLADARVCAIMMLILLOQEALEKLVILHAASASCNGLFYFVIAW	780
QY	781	YIKGRVVEPLATYSLTGMSFSLLILALPOQAYAVDASVHGOIGALLVMITLFTLTPRGYK	840
Db	781	YIKGRVVEPLATYSLTGMSFSLLILALPOQAYAVDASVHGOIGALLVMITLFTLTPRGYK	840
QY	841	TLLSRFLMWCYLLTLGEBAMQOBAMPQVAGGDDIIMAVAITPBGVPIITKMLAVL	900
Db	841	TLLSRFLMWCYLLTLGEBAMQOBAMPQVAGGDDIIMAVAITPBGVPIITKMLAVL	900
QY	901	GPAYLLKCALTRVPYFVRAHALLRMCTMARHLAAGRYVOMALLALGRWGTYYIDHLTPM	960
Db	901	GPAYLLKCALTRVPYFVRAHALLRMCTMARHLAAGRYVOMALLALGRWGTYYIDHLTPM	960
QY	961	SDMAASGLRLAVALVEPIIFSPMEKVI VMGAETPAACGDIILHGLPVSARLREVLGAD	1020
Db	961	SDMAASGLRLAVALVEPIIFSPMEKVI VMGAETPAACGDIILHGLPVSARLREVLGAD	1020
QY	1021	GYTSKMSLLAPITAYAOITRGLLGTIVSMTGRDKTEQAGEIOVLSTVTSPLGTSISG	1080
Db	1021	GYTSKMSLLAPITAYAOITRGLLGTIVSMTGRDKTEQAGEIOVLSTVTSPLGTSISG	1080
QY	1081	VLMWYVYHGAAGKTLAAGSRGPTVQWYSSAEGDLVGMPSPPGTKSLEPCTCGAVDLXYLVTRN	1140
Db	1081	VLMWYVYHGAAGKTLAAGSRGPTVQWYSSAEGDLVGMPSPPGTKSLEPCTCGAVDLXYLVTRN	1140
QY	1141	ADVPARRRGDKRGALLSPRLSTLKSGSGPVLCPRGHAGVFPAAVCSRGVAKSIDFI	1200
Db	1141	ADVPARRRGDKRGALLSPRLSTLKSGSGPVLCPRGHAGVFPAAVCSRGVAKSIDFI	1200

Db 510 ----- 509
Qy 1201 PVEITLDIVTRSPPTSDNSTPPPAVPOTVQVGLHAFTSGSKSTKVPVAAQGYKVLVLP 1260
Db 510 ----- 509
Qy 1261 SVAATLGFGLSKAHGINPNI RTGVRTVTGAPITTSYGKFLADGGCAGAGAYDIITCD 1320
Db 510 ----- 509
Qy 1321 ECHAVDSPTIIGITVLDOAETAGVRLTVLATATPPGSVTPPHPIEEVALGQBEIPEY 1380
Db 510 ----- 509
Qy 1381 GRALPLSTIKGGRHLIFCHSKKKCDELAALRGMLNSVAYYRGLDVSIPTQGVVVA 1440
Db 510 ----- 509
Qy 1441 TDALMTGYTGPDSVIDCNVAVTQVVDPSLDPTFTITQIVQDAVSRSGRRGTRGRRL 1500
Db 510 ----- 509
Qy 1501 GIYRVSTGERASGMFDSVLC ECYDAGAWELTPESETVRLRAYFNTPGLPVCQDHL 1560
Db 510 ----- 509
Qy 1561 FWEAVFTGLTHI DAHFLSQTKQSGENFAYLTAYQATVCARAKAPPSMDVMKCLTRLKP 1620
Db 510 ----- 509
Qy 1621 TLVGPRTLRLSGVSTNEVTILTHPYTKYIATCQADLEVMSTWVLGGVLAANAAYCLA 1680
Db 510 ----- 509
Qy 1681 TGCVCIIIGRLHINORAVVAPDKEVLYEAFDEMECASRAALIEBQRIAEMLKSKIOGL 1740
Db 510 ----- 509
Qy 1741 QOASKOADIQPTQVQASWPKEQFWAKMMNFI SGIQYLAGLSTLPGNPAVASMMAFSNA 1800
Db 510 ----- 509
Qy 1801 LTSPLSTSTTLLNLTGLGMLASQIAPAGATGFVSVGLGA VSGISGLKVLVDLLAGYG 1860
Db 510 ----- 509
Qy 1861 AGISGALVAFKIMSGEKPSMEDVVMILPGLSPGALVVGICAILRRHVGPBGAVQMM 1920
Db 510 ----- 509
Qy 1921 NRLIAFASRGHNAFTHYTESDASQAVTQLIGSLTTSLLRLHNMWITEDCPIPCGGSW 1980
Db 510 ----- 509
Qy 1981 LRVDWDMVCTIITDPFKMVLTKLFPKMPGLPFPVSCQKGYKGMAGTGIMTRCPGAGNIS 2040
Db 510 ----- 509
Qy 2041 GNVRLGSRWRTGPKTCNNIWOGTFPINCYTGQCVKPAKVAIMWEAASEVAEVTQH 2100
Db 510 ----- 509
Qy 2101 GSYHYITGLTDDNLKVPQOLPSPEFFSWDGVQIHRFAPTPKPFREDEVSCVGLNSFV 2160
Db 510 ----- 509
Qy 2161 GSQLPCEDEPPTDVLMMLTDSHITTAETAARLARSGSPPSRASSASQLSAPSLATCT 2220
Db 510 ----- 509
Qy 2221 THGKAYDVDMVDANLFMGGDVTRIESGSKVVVLDLDPWVEERSDLBPSIPSEYMLPKKR 2280

Db 510 ----- 509
Qy 2281 FPPALPAMARPDYNPPLVESMKRPDYOPATVAGCALPPPKRTPTPPRRRRRTVGLSEDSI 2340
Db 510 ----- 509
Qy 2341 GDALQOLAIKSGQPPPSGDSGLSTGAGADSGSQTPPELALSETGISMPLEGLG 2400
Db 510 ----- 509
Qy 2401 DPDLPEQVEPQPPPOGGVAAPGSDSGSWSTCSEBDSVCCSMSYMTGALITCSPBE 2460
Db 510 ----- 509
Qy 2461 EKLPIINPSNLSLRHNKVCTTTSASLBAKAVTFDMQVLDSSYYDSVLKIKLAASKV 2520
Db 510 ----- 509
Qy 2521 TARTLTWEACQLTPPHSARSKYGFKAKEVRSLSGRAVNHIKSVKDLLEDESEPIPTTI 2580
Db 510 ----- 509
Qy 2581 MAKNEVFCVDPPTKGKKARLIVYPDLGVRVCEKALYDTQKLPQAVMGASVGFQYSPA 2640
Db 510 ----- 509
Qy 2641 QRVFLLKAMAEKXDPMGFSYDTRCFDSTVTBRDITREESTYRACSLPEAHTAHSLTE 2700
Db 510 ----- 541
Qy 2701 RLTVGPMFNKSGQTCGYRCRASGVLTTSMGNTITCYVKALAAKAGIIPMLVCGD 2760
Db 542 RLTVGPMFNKSGQTCGYRCRASGVLTTSMGNTITCYVKALAAKAGIIPMLVCGD 601
Qy 2761 DLVVISQGTBEDERNIPATEAMTRYSDPPDPREYDLELITSCSSNVSVAGPQG 2820
Db 602 DLVVISQGTBEDERNIPATEAMTRYSDPPDPREYDLELITSCSSNVSVAGPQG 661
Qy 2821 RRRYLTPTDPTTPITARAAMETVRRHS PVNSWGNIIQVAPTIWAVMLTHPFSILMAODT 2880
Db 662 RRRYLTPTDPTTPITARAAMETVRRHS PVNSWGNIIQVAPTIWAVMLTHPFSILMAODT 721
Qy 2881 LDONLNFMYGAVSVSPDLDPATIERLHGDAFSLHTYTPHELTRVASALRKLGAAPLR 2940
Db 722 LDONLNFMYGAVSVSPDLDPATIERLHGDAFSLHTYTPHELTRVASALRKLGAAPLR 781
Qy 2941 AMKSRARA VRA SLISRGRAAVCGRYLFNMAVTKLKLTPLEARLLDLSSWFTVAGCGG 3000
Db 782 AMKSRARA VRA SLISRGRAAVCGRYLFNMAVTKLKLTPLEARLLDLSSWFTVAGCGG 841
Qy 3001 DIHVSRAARPRLLLFGILLFPVGLFLPAR 3033
Db 842 DIHVSRAARPRLLLFGILLFPVGLFLPAR 874

RESULT 13
PC2219
polypeptide - hepatitis C virus (type 5a) (fragments)
N:Contains: core protein; E1 (carboxyl end); E2/NS1 (amino end); NS3 protein; NS4A prote
C:Species: hepatitis C virus
C:date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: PC2219
R:Stuyver, J.; Arnhem, W.V.; Wyseur, A.; Maertens, G.
Biochem. Biophys. Res. Commun. 202, 1308-1314, 1994
A:title: Cloning and phylogenetic analysis of the core, E2, and NS3/NS4 regions of the
A:Reference number: PC2219; MUID:94338342; PMID:7520237
A:Accession: PC2219
A:Molecule type: mRNA
A:Residues: 1-876 <STU>
A:Cross-References: UNIPROT:081242; GB:L29577; GB:L29578; GB:L29579
A:Experimental source: serum
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: glycoprotein

F:1-191/Product: core #status predicted <COB>
 F:68-78/Region: variable
 F:192-247/Product: E1 (carboxyl end) #status predicted <ERB>
 F:248-411/Product: E2/NS1 (amino end) #status predicted <ENR>
 F:248-338/Region: E2
 F:339-411/Region: NS1 (amino end)
 F:412-783/Product: NS3 #status predicted <NSR>
 F:784-837/Product: NS4 #status predicted <NSA>
 F:838-876/Product: NS4B #status predicted <NSB>
 F:281,287,294,312,340/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 20.2%; Score 3263; DB 2; Length 876;
 Best Local Similarity 38.1%; Pred. No. 1.8e-178;

Matches 667; Conservative 86; Mismatches 121; Indels 878; Gaps 3;

```

QY 1 MSTNPRQRTKNTNRPRDVKFPGGGQIVGGVYLPRRPRPLGRATRTKTSERSQPRG 60
DB 1 MSTNPRQRTKNTNRPRDVKFPGGGQIVGGVYLPRRPRPMGVRATRTKTSERSQPRG 60
QY 61 RRQPIKRRSTGSKMGKPGYPMPLYNBGLAGWMLSPGSRPSMGPNDRHRSRVNG 120
DB 61 RRQPIKRRSTGSKMGKPGYPMPLYNBGLAGWMLSPGSRPSMGPNDRHRSRVNG 120
QY 121 KYVDLTTCGRADMGITPVVGAFLGGVAAALAGVVLLEDGVNFATGNIPEGCSFSLILA 180
DB 121 KYVDLTTCGRADMGITPVVGAFLGGVAAALAGVVLLEDGVNFATGNIPEGCSFSLILA 180
QY 181 LLSCLITPVSAAEVKNIITGYVWNTDCTNDSITWQAAVLHVPGCVCEKYGNASQCWI 240
DB 181 LLSCLITPVSAAEVKNIITGYVWNTDCTNDSITWQAAVLHVPGCVCEKYGNASQCWI 240
QY 241 PVSNNVAVORPGALTOGLRTHIDMVVMSATLCSALYVGDLCGGMLAQMFIVSPQHWF 300
DB 241 PVSNNVAVORPGALTOGLRTHIDMVVMSATLCSALYVGDLCGGMLAQMFIVSPQHWF 300
QY 192 ----- 191
DB 192 ----- 191
QY 301 VDDCNGSIYPGITTHGRMAMMMNMPATMTLAAVREVIIDISGAHMGVFGIA 360
DB 301 VDDCNGSIYPGITTHGRMAMMMNMPATMTLAAVREVIIDISGAHMGVFGIA 360
QY 192 ----- 224
DB 192 ----- 224
QY 361 YFSGGAMAKVVVILLAAAGVDAARTHTVGSAAQTGRLTSLFDMPROKIQLVNTNGSM 420
DB 361 YFSGGAMAKVVVILLAAAGVDAARTHTVGSAAQTGRLTSLFDMPROKIQLVNTNGSM 420
QY 225 YVASVANNKVVVLLFPAQVDATQISGSSNQTTGYIASFTTRAAQKQLQIINTNGSM 284
DB 225 YVASVANNKVVVLLFPAQVDATQISGSSNQTTGYIASFTTRAAQKQLQIINTNGSM 284
QY 421 HINRTALNCNDSLHTGFIASLFTYHSFNSSGCEPERSACRSLEAFVNGALQYEDNTN 480
DB 421 HINRTALNCNDSLHTGFIASLFTYHSFNSSGCEPERSACRSLEAFVNGALQYEDNTN 480
QY 285 HINRTALNCNDSLQGTGFIAGLFYHAFKFNSSGCDPRMASCRAALFFDQGWGITISYA-NISG 343
DB 285 HINRTALNCNDSLQGTGFIAGLFYHAFKFNSSGCDPRMASCRAALFFDQGWGITISYA-NISG 343
QY 481 PEDMRPYCMHYPRGCGVVSAAKTVCGPVYCFTPSPVVGTTDRLAGPYTWGENEDVVL 540
DB 481 PEDMRPYCMHYPRGCGVVSAAKTVCGPVYCFTPSPVVGTTDRLAGPYTWGENEDVVL 540
QY 344 PSDDKYCMHYPRPCGVVPAQVCGPVYCFTPSPVVGTTDSKGHTYNGSNVTDFFL 403
DB 344 PSDDKYCMHYPRPCGVVPAQVCGPVYCFTPSPVVGTTDSKGHTYNGSNVTDFFL 403
QY 541 LNSTREPLGSMFCGTMMNSSGYTKTGAPPCRTRADFNASTDILCTDCEFRKHPDTTYLK 600
DB 541 LNSTREPLGSMFCGTMMNSSGYTKTGAPPCRTRADFNASTDILCTDCEFRKHPDTTYLK 600
QY 404 MNNTTRP----- 410
DB 404 MNNTTRP----- 410
QY 601 CGSGPWLTPRCLIDPYRLMHYPCVNYTIFKIRMYGVGEHRLTAACNFTGRDRCNLED 660
DB 601 CGSGPWLTPRCLIDPYRLMHYPCVNYTIFKIRMYGVGEHRLTAACNFTGRDRCNLED 660
QY 411 ----- 410
DB 411 ----- 410
QY 661 RDRSQSLPLHSTTEWAILPCSYSDLPALSTGLLHONIVDVQPMYGLSPALTKYIVRW 720
DB 661 RDRSQSLPLHSTTEWAILPCSYSDLPALSTGLLHONIVDVQPMYGLSPALTKYIVRW 720
QY 411 ----- 410
DB 411 ----- 410
QY 721 EWVILLFLIADARVACLMMLILLGOAEALAEKVLILHAASASNGFLYFVIFVAAW 780
DB 721 EWVILLFLIADARVACLMMLILLGOAEALAEKVLILHAASASNGFLYFVIFVAAW 780
QY 411 ----- 410
DB 411 ----- 410
QY 781 YIKGRVPLATYSLTGLMSFSLILLALPOQAVYDASVHGIGALLVMITLFTLTPGYK 840
DB 781 YIKGRVPLATYSLTGLMSFSLILLALPOQAVYDASVHGIGALLVMITLFTLTPGYK 840
QY 411 ----- 410
DB 411 ----- 410
QY 841 TLISRLMWMLCYLLTLGEMVQGEWAPPMQVGRGRDGIWAVALFYPGVVFDTIKMLLAVL 900
DB 841 TLISRLMWMLCYLLTLGEMVQGEWAPPMQVGRGRDGIWAVALFYPGVVFDTIKMLLAVL 900

```

```

DB 411 ----- 410
QY 901 GPAYLLKGLATRVFYFVRAHALRLMCTWARIAGGRVYQMALLAGRWGTGYIYDHLTPM 960
DB 411 ----- 410
QY 961 SDMAASGLRLDAVAVERPIIFSPMEKKVILWGAETACGDIILHGLPVSARLREVLGPA 1020
DB 411 ----- 410
QY 1021 GYTSKMSGLAPITRAYAQTGGLGTIVSMTRGRDKTEQAGEIQVLSVTQSFLGTSISG 1080
DB 411 ----- 410
QY 1081 VLMTVYHGAGNKTLAGSRGPVTQWYSABGDLVGMPSBPTKSLPCTCGAVDLYLVTRN 1140
DB 411 ----- 410
QY 1141 ADVIIPARRRGDKRGALLSPRLSTLKGSSGGPVLCPRGAVGVPRRAVCSRGVAKSIDFI 1200
DB 411 ----- 410
QY 1201 PVEITDIYTRBPTSDNSTPRAVPOTYGVYLIHAPTSKSKTKVPVAAAGYKVLVLP 1260
DB 411 ----- 410
QY 1261 SVAATLFGAYLSKAGINPNIIRGVRTVTTGADITVSTYSGFLADGGCAGGAYDIICD 1320
DB 411 ----- 1320
QY 1321 ECHAVDSITTIIGITVLDQAEETAGVRLTVLATATPQSVTTPHNIEVALGQGEIIPFY 1380
DB 1321 ECHAVDSITTIIGITVLDQAEETAGVRLTVLATATPQSVTTPHNIEVALGQGEIIPFY 1380
QY 443 ECHSQDATTIIGITVLDQAEETAGVRLTVLATATPQSVTTPHNIEVALGQGEIIPFY 502
DB 443 ECHSQDATTIIGITVLDQAEETAGVRLTVLATATPQSVTTPHNIEVALGQGEIIPFY 502
QY 1381 GRATPLSYIKGGRHLIFCHSKKKCDELAAALRGGLNSVAYYRGHLDVSIPTQGDVVVA 1440
DB 1381 GRATPLSYIKGGRHLIFCHSKKKCDELAAALRGGLNSVAYYRGHLDVSIPTQGDVVVA 1440
QY 503 GRATPLSYIKGGRHLIFCHSKKKCDELAAALRGGLNSVAYYRGHLDVSIPTQGDVVVA 562
DB 503 GRATPLSYIKGGRHLIFCHSKKKCDELAAALRGGLNSVAYYRGHLDVSIPTQGDVVVA 562
QY 1441 TDALMTGYTGDFDSYIDCNVAVTQVVDPSLPTFTITTIQIYPODAVNSRSGRGRGRGL 1500
DB 1441 TDALMTGYTGDFDSYIDCNVAVTQVVDPSLPTFTITTIQIYPODAVNSRSGRGRGRGL 1500
QY 563 TDALMTGYTGDFDSYIDCNVAVTQVVDPSLPTFTITTIQIYPODAVNSRSGRGRGRGL 622
DB 563 TDALMTGYTGDFDSYIDCNVAVTQVVDPSLPTFTITTIQIYPODAVNSRSGRGRGRGL 622
QY 1501 GIYRYVSTGERASGMFDSVULCECYDAGANAYELTPSETTYRLAAYFNTPLPYCQDHL 1560
DB 1501 GIYRYVSTGERASGMFDSVULCECYDAGANAYELTPSETTYRLAAYFNTPLPYCQDHL 1560
QY 623 GIYRYVSTGERASGMFDSVULCECYDAGANAYELTPSETTYRLAAYFNTPLPYCQDHL 682
DB 623 GIYRYVSTGERASGMFDSVULCECYDAGANAYELTPSETTYRLAAYFNTPLPYCQDHL 682
QY 1561 FMEAVFTGLTHIDAFHSQTKOSGENFAYLTAAYQATVCARAKAPPSMDVMKCLTRLKP 1620
DB 1561 FMEAVFTGLTHIDAFHSQTKOSGENFAYLTAAYQATVCARAKAPPSMDVMKCLTRLKP 1620
QY 663 FMEAVFTGLTHIDAFHSQTKOSGENFAYLTAAYQATVCARAKAPPSMDVMKCLTRLKP 742
DB 663 FMEAVFTGLTHIDAFHSQTKOSGENFAYLTAAYQATVCARAKAPPSMDVMKCLTRLKP 742
QY 1621 TLVGFPTPLLYRLGSVTNEVTLTHPVTKYIATCMQADLEVMSTWVLAGVLAAYVCLA 1680
DB 1621 TLVGFPTPLLYRLGSVTNEVTLTHPVTKYIATCMQADLEVMSTWVLAGVLAAYVCLA 1680
QY 743 TLVGFPTPLLYRLGSVTNEVTLTHPVTKYIATCMQADLEVMSTWVLAGVLAAYVCLA 802
DB 743 TLVGFPTPLLYRLGSVTNEVTLTHPVTKYIATCMQADLEVMSTWVLAGVLAAYVCLA 802
QY 1681 TGCVCITIRLHINORAVAVAPKEVLYEAFDEMEBCASALIBEGQRIAEMLKSIQGL 1740
DB 1681 TGCVCITIRLHINORAVAVAPKEVLYEAFDEMEBCASALIBEGQRIAEMLKSIQGL 1740
QY 803 VGSVAIVRRIILSGKPAIIPREALYQOFDEMEBCASALPYMDETRIALAGFKKVLGFI 862
DB 803 VGSVAIVRRIILSGKPAIIPREALYQOFDEMEBCASALPYMDETRIALAGFKKVLGFI 862
QY 1741 QQASKOADIOP 1752
DB 1741 QQASKOADIOP 1752
QY 863 STTGQKAEITLKP 874
DB 863 STTGQKAEITLKP 874

```

RESULT 14

S19875
 genome polyprotein - hepatitis C virus (isolate JK3) (fragment)
 N/contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein
 C/species: hepatitis C virus
 A/variety: isolate JK3
 C/date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
 C/accession: S19875
 R/Honda, M.; Kaneko, S.; Maeshi, U.; Kobayashi, K.; Murakami, S.
 submitted to the EMBL Data Library, September 1991

A:Description: Sequence analysis of putative structural regions of Hepatitis C Virus 1sc
A:Reference number: S18029
A:Accession: S19875
A:Molecule type: genomic RNA
A:Residues: 1-782 <HON>
A:Cross-references: UNIPROT:068951; EMBL:X61592; NID:G59482; PID:CAA43789.1; PID:G59483
A:Experimental source: isolate JK3
A:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural
F:1-191/Product: core protein #status predicted <Mat1>
F:192-383/Product: envelope protein 1 #status predicted <Mat2>
F:384-733/Product: NS1/E2 protein #status predicted <Mat3>
F:734-782/Product: nonstructural protein 2 (fragment) #status predicted <Mat4>

Query Match 20.0%; Score 3231; DB 2; Length 782;
Best Local Similarity 70.7%; Pred. No. 1e-176;
Matches 556; Conservative 104; Mismatches 122; Indels 4; Gaps 2;

QY 1 MSTNKPORKTKRNTNRPRPDVKFPGGQIVGVYLLRRGRRLGVRATRKTSERQPRG 60
DB 1 MSTNKPORKTKRNTNRPRPDVKFPGGQIVGVYLLRRGRRLGVRATRKTSERQPRG 60
QY 61 RRPIPKDRSTGKSWGKPGYWPPLYGNEGIGMACWLLSPRGRSPSWGPNDRHRSRVNG 120
DB 61 RRPIPKAROBEGRAMAOPGYWPPLYGNEGIGMTGMLSPRGRSPSWGPTDRRRSRNLG 120
QY 121 KYIDLTGSPADLMQYIVVGAPLGGVAPALAHGVRVLEDGNFATGNLPGCSFSIFLLA 180
DB 121 KYIDLTGSPADLMQYIVVGAPLGGVAPALAHGVRVLEDGNVATGNLPGCSFSIFLLA 180
QY 181 LLSCTTTPVSAAEVKNISTGYWVTNDCTNDSITWLOQAALVHPGCVCEKVGANSQCWI 240
DB 181 LLSCTTTPVSAAEVKNISTGYWVTNDCTNDSITWLOQAALVHPGCVCEKVGANSQCWI 240
QY 241 PVSPNVAVORPGALTOGLRTHIDMYVMSATLCSALYVDLCGGVMLAAQMFIVSQHMF 300
DB 241 ALTLTLAARNSSIPFTTIRRHVDLLVGAALCSAMYVDLCGSVFLVSQLETFPSRRYET 300
QY 301 VQDCNCSLYPGTITGHRAMDMMMNSPTALVAVSOLLRIPQAVDVMAAGAMVLAGLA 360
DB 301 VQDCNCSLYPGTITGHRAMDMMMNSPTALVAVSOLLRIPQAVDVMAAGAMVLAGLA 360
QY 361 YFSMGAQAKVIVVILLAGVDARTHTVGSAAQTGRLTSLFDMPROKIQLVNTNGSM 420
DB 361 YFSMGAQAKVIVVILLAGVDARTHTVGSAAQTGRLTSLFDMPROKIQLVNTNGSM 420
QY 421 HINRTALNCNDSLHTGFIASLEFYTHSFNSGCCPERMSACRSIEAFVGMALQYEDNTN 480
DB 421 HINRTALNCNDSLHTGFIASLEFYTHSFNSGCCPERMSACRSIEAFVGMALQYEDNTN 480
QY 481 PEDMRPYCMHYPRRCGVSAKTVCGPRVYCFPPSPVVVGTDRGLCAPTYTWGENETDVL 540
DB 479 TTDOKPYCMHYAPRCGIVPASQVCGPYVCFPPSPVVVGTDRGLCAPTYTWGENETDVL 538
QY 541 INSTRPPLGSGWFGCTWMMNSGGYTKTCGAPPCRTADFNASTDLCPTCCFRGHPPTTYLK 600
DB 539 LNNRPPGGMNFGCTWMMNSGGYTKTCGAPPCRTADFNASTDLCPTCCFRGHPPTTYLK 596
QY 601 CGSGFWLTPRCLIDYPRYLAMYPCVTYNTIFKIRMYVGVGVEHRLTAACNFTGRDCNLED 660
DB 597 CGSGFWLTPRCLIDYPRYLAMYPCVTYNTIFKIRMYVGVGVEHRLTAACNFTGRDCNLED 656
QY 661 RDRSGLSLHSTTEMWALLPGYSYDLPALSTGLHLHONIVDVOYMGSLPALTKYIRW 720
DB 657 RDRSGLSLHSTTEMWALLPGYSYDLPALSTGLHLHONIVDVOYMGSLPALTKYIRW 716
QY 721 EMWILLFLLDARVACIMMLILGOAALAEKLVIIHAASAASCNGLFVIFVAVAM 780
DB 717 EYILLFLLDARVACIMMLILGOAALAEKLVIIHAASAASCNGLFVIFVAVAM 776
QY 781 YIKGRV 786
DB 777 YIKGRV 782

RESULT 15
S19876
genome polypeptide - hepatitis C virus (isolate JK3) (fragment)
N:Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein
C:Species: hepatitis C virus
A:Variety: isolate JK5
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: S19876
R:Honda, M.; Kaneko, S.; Maeshi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A:Description: Sequence analysis of putative structural regions of Hepatitis C Virus 1sc
A:Reference number: S18029
A:Accession: S19876
A:Molecule type: genomic RNA
A:Residues: 1-782 <HON>
A:Cross-references: UNIPROT:068953; EMBL:X61595; NID:G59486; PID:CAA43792.1; PID:G59487
A:Experimental source: isolate JK5
A:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural
F:1-191/Product: core protein #status predicted <Mat1>
F:192-383/Product: envelope protein 1 #status predicted <Mat2>
F:384-733/Product: NS1/E2 protein #status predicted <Mat3>
F:734-782/Product: nonstructural protein 2 (fragment) #status predicted <Mat4>

Query Match 19.9%; Score 3214; DB 2; Length 782;
Best Local Similarity 70.6%; Pred. No. 9.5e-176;
Matches 555; Conservative 102; Mismatches 125; Indels 4; Gaps 2;

QY 1 MSTNKPORKTKRNTNRPRPDVKFPGGQIVGVYLLRRGRRLGVRATRKTSERQPRG 60
DB 1 MSTNKPORKTKRNTNRPRPDVKFPGGQIVGVYLLRRGRRLGVRATRKTSERQPRG 60
QY 61 RRPIPKDRSTGKSWGKPGYWPPLYGNEGIGMACWLLSPRGRSPSWGPNDRHRSRVNG 120
DB 61 RRPIPKAROBEGRAMAOPGYWPPLYGNEGIGMACWLLSPRGRSPSWGPTDRRRSRNLG 120
QY 121 KYIDLTGSPADLMQYIVVGAPLGGVAPALAHGVRVLEDGNFATGNLPGCSFSIFLLA 180
DB 121 KYIDLTGSPADLMQYIVVGAPLGGVAPALAHGVRVLEDGNVATGNLPGCSFSIFLLA 180
QY 181 LLSCTTTPVSAAEVKNISTGYWVTNDCTNDSITWLOQAALVHPGCVCEKVGANSQCWI 240
DB 181 LLSCTTTPVSAAEVKNISTGYWVTNDCTNDSITWLOQAALVHPGCVCEKVGANSQCWI 240
QY 241 PVSPNVAVORPGALTOGLRTHIDMYVMSATLCSALYVDLCGGVMLAAQMFIVSQHMF 300
DB 241 ALTLTLAARNSSIPFTTIRRHVDLLVGAALCSAMYVDLCGSVFLVSQLETFPSRRYET 300
QY 301 VQDCNCSLYPGTITGHRAMDMMMNSPTALVAVSOLLRIPQAVDVMAAGAMVLAGLA 360
DB 301 VQDCNCSLYPGTITGHRAMDMMMNSPTALVAVSOLLRIPQAVDVMAAGAMVLAGLA 360
QY 361 YFSMGAQAKVIVVILLAGVDARTHTVGSAAQTGRLTSLFDMPROKIQLVNTNGSM 420
DB 361 YFSMGAQAKVIVVILLAGVDARTHTVGSAAQTGRLTSLFDMPROKIQLVNTNGSM 420
QY 421 HINRTALNCNDSLHTGFIASLEFYTHSFNSGCCPERMSACRSIEAFVGMALQYEDNTN 480
DB 421 HINRTALNCNDSLHTGFIASLEFYTHSFNSGCCPERMSACRSIEAFVGMALQYEDNTN 480
QY 481 PEDMRPYCMHYPRRCGVSAKTVCGPRVYCFPPSPVVVGTDRGLCAPTYTWGENETDVL 540
DB 479 TTDOKPYCMHYAPRCGIVPASQVCGPYVCFPPSPVVVGTDRGLCAPTYTWGENETDVL 538
QY 541 INSTRPPLGSGWFGCTWMMNSGGYTKTCGAPPCRTADFNASTDLCPTCCFRGHPPTTYLK 600
DB 539 LNNRPPGGMNFGCTWMMNSGGYTKTCGAPPCRTADFNASTDLCPTCCFRGHPPTTYLK 596
QY 601 CGSGFWLTPRCLIDYPRYLAMYPCVTYNTIFKIRMYVGVGVEHRLTAACNFTGRDCNLED 660
DB 597 CGSGFWLTPRCLIDYPRYLAMYPCVTYNTIFKIRMYVGVGVEHRLTAACNFTGRDCNLED 656

Qy	661	RDRSQSPPLHSTTEWAILPCSYSDLPALSTGHLHONIVDVQFMVYGLSPALTKYVRW	720
Db	657	RDRSELSPILRATTEMQILPCSFITLPALSTGHLHONIVDVQYLYGSAVVSIVIKW	716
Qy	721	EWVILFLLIADARVCACIMMLILGQAEALEKVIYIHAASASCNGFLYFVIFVAAW	780
Db	717	EYVLLFLLIADARVCACIMMMLIAQAEALENLVVINAASVAGAHGILSFLVFCAAW	776
Qy	781	YIKGRV	786
Db	777	YIKGRL	782

Search completed: October 28, 2005, 16:06:03
Job time : 97 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2005, 15:36:16 ; Search time 328 Seconds

(without alignments)
4735.171 Million cell updates/sec

Title: US-09-980-559-2

Perfect score: 16120

Sequence: 1 MSTNPKPQKRTKMTNRKPPQ.....LLFGLLLFVGVLFLPAR 3033

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16120	100.0	3033	2	Q9QF35
2	15830	98.2	3033	1	P26660 h genome po
3	15426	95.7	3033	2	Q9Q9A7
4	15394	95.5	3033	2	Q9Q9B6
5	15356	95.3	3033	2	Q91ZAA
6	15355	95.3	3033	2	Q91ZAA
7	15354	95.2	3033	2	Q91ZAA
8	15350	95.2	3033	2	Q91ZAA
9	15347	95.2	3033	2	Q91ZAA
10	15324	95.1	3033	2	Q91ZAA
11	15322	95.0	3033	2	Q91ZAA
12	15321	95.0	3033	2	Q91ZAA
13	15311	95.0	3033	2	Q9Q9B0
14	15266	94.7	3033	2	Q9Q9A8
15	15249	94.6	3033	2	Q9Q9A9
16	15207	94.3	3033	2	Q9Q9A8
17	15115	93.8	3032	2	Q9Q9A8
18	14943	92.7	3033	2	Q9Q9A8
19	14936	89.1	3037	2	Q68749
20	14185	88.0	3033	2	Q9Q9A1
21	14014	86.9	3033	2	Q9Q9A1
22	14013	86.9	3033	2	Q9Q9A1
23	14007	86.9	3033	2	Q9Q9A1
24	14005	86.9	3033	2	Q9Q9A1
25	14003	86.9	3033	2	Q9Q9A1
26	13998	86.8	3033	2	Q9Q9A1
27	13990	86.8	3033	2	Q9Q9A1
28	13987	86.8	3033	2	Q9Q9A1
29	13987	86.8	3033	2	Q9Q9A1
30	13976	86.7	3033	2	Q9Q9A1
31	13971	86.7	3033	1	P26661 h genome po

32	13965	86.6	3033	2	Q9Q9A1
33	13961	86.6	3033	2	Q9Q9A1
34	13960	86.6	3033	2	Q9Q9A1
35	13939	86.5	3033	2	Q9Q9A1
36	13937	86.5	3033	2	Q9Q9A1
37	13919	86.3	3033	2	Q9Q9A1
38	13917	86.3	3033	2	Q9Q9A1
39	13914	86.3	3033	2	Q9Q9A1
40	13912	86.3	3033	2	Q9Q9A1
41	13885	86.1	3033	2	Q9Q9A1
42	13884	86.1	3033	2	Q9Q9A1
43	13884	86.1	3033	2	Q9Q9A1
44	13049	80.9	3015	2	Q9Q9A1
45	12925	80.2	3015	2	Q9Q9A1

ALIGNMENTS

RESULT 1
Q9QF35 PRELIMINARY: PRT: 3033 AA.
AC Q9QF35;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage, Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HC-36CH; PubMed=10489358; DOI=10.1006/viro.1999.9889;
RX MEDLINE=99420396; PubMed=10489358; DOI=10.1006/viro.1999.9889;
RA Yanagi M., Purcell R.H., Emerson S.U., Bukh J.;
RT "Hepatitis C virus: an infectious molecular clone of a second major
RT genotype (2a) and lack of viability of intertypic 1a and 2a
RT chimeras".
RL Virology 262:250-263(1999).
DR EMBL, AF177036; AAF01178.1; -.
DR HSP; O8JYS1; ICWX.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; F:transcription; IEA.
DR GO; GO:0019079; F:viral genome replication; IEA.
DR GO; GO:0019087; F:viral transformation; IEA.
DR InterPro: IPR000345; CytoC_heme_BS.
DR InterPro: IPR001410; DEAD/DEAH_BS.
DR InterPro: IPR011545; DEAD/DEAH_N.
DR InterPro: IPR002522; HCV_capsid.
DR InterPro: IPR002521; HCV_core.
DR InterPro: IPR002519; HCV_env.
DR InterPro: IPR002531; HCV_NS1.
DR InterPro: IPR002745; HCV_NS4a.
DR InterPro: IPR001490; HCV_NS4b.
DR InterPro: IPR002868; HCV_NS5a.
DR InterPro: IPR002166; HCV_NS5a.
DR InterPro: IPR004109; Peptidase_S29.
DR InterPro: IPR009003; Peptidase_S29.
DR InterPro: IPR002518; Peptidase_S29.
DR InterPro: IPR002129; Pyridoxal_dec.
DR InterPro: IPR007095; RNA_pol_DS_P5.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.

DR	Pfam; PF01542; HCV_core; 1.	
DR	Pfam; PF01539; HCV_env; 1.	
DR	Pfam; PF01560; HCV_NS1; 1.	
DR	Pfam; PF01538; HCV_NS2; 1.	
DR	Pfam; PF02907; HCV_NS3; 1.	
DR	Pfam; PF01006; HCV_NS4; 1.	
DR	Pfam; PF01001; HCV_NS4b; 1.	
DR	Pfam; PF01506; HCV_NS5a; 1.	
DR	Pfam; PF00998; Viral_Rdrp; 1.	
DR	SMART; SM00487; DEXDC; 1.	
DR	PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.	
DR	PROSITE; PS00392; DDC_GAD_HDC_YDC; UNKNOWN_1.	
KW	Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;	
KW	Polypeptide; Transmembrane.	
SEQUENCE	3033 AA; 329264 MW; 6D894082FF685C7 CRC64;	
Query Match	100.0%; Score 16120; DB 2; Length 3033;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 3033; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MSTNPKRQRTKNTNRBRPODYKFPGGGQIVGGVYLLPRGRPLGVRATRKTSERSQPRG	60
DB	1 MSTNPKRQRTKNTNRBRPODYKFPGGGQIVGGVYLLPRGRPLGVRATRKTSERSQPRG	60
QY	61 RRQPIFKDRRSTGSKGKPGYPMPLYGNBGLGNAGWLLSPRGRPSWGNDRPHRSRNVG	120
DB	61 RRQPIFKDRRSTGSKGKPGYPMPLYGNBGLGNAGWLLSPRGRPSWGNDRPHRSRNVG	120
QY	121 KVIDTTLTCGPADLMGTYIPVVGARLGGVARRLHGVNLEEDGVNFALGNIPGCSFSIFLLA	180
DB	121 KVIDTTLTCGPADLMGTYIPVVGARLGGVARRLHGVNLEEDGVNFALGNIPGCSFSIFLLA	180
QY	181 LLSCTITPVSAAEVKINISTGYMTNDCTNDSITWOLAAVLAHPGCPCEKYGANASQCI	240
DB	181 LLSCTITPVSAAEVKINISTGYMTNDCTNDSITWOLAAVLAHPGCPCEKYGANASQCI	240
QY	241 PVSNNVAVORPGLTGLRTHIDMVVMSATLCSALYVGDLCGGVLAQMFIIVSPQHMF	300
DB	241 PVSNNVAVORPGLTGLRTHIDMVVMSATLCSALYVGDLCGGVLAQMFIIVSPQHMF	300
QY	301 VQDNCNSIYRGITTTGRMADMMWMSPTATMLAYAMRVEYITIDIIISGAHGVNFGLA	360
DB	301 VQDNCNSIYRGITTTGRMADMMWMSPTATMLAYAMRVEYITIDIIISGAHGVNFGLA	360
QY	361 YFSMGGAMAKVVYILLAAGVDAARTHTVGGSAQTGRLTSLFDMGPROKIQLVNTNGSM	420
DB	361 YFSMGGAMAKVVYILLAAGVDAARTHTVGGSAQTGRLTSLFDMGPROKIQLVNTNGSM	420
QY	421 HINRTALNCNDSLHTGFIASLFTYHSFNSSGCEPERSACSIIIAFVVGALQYEDNVN	480
DB	421 HINRTALNCNDSLHTGFIASLFTYHSFNSSGCEPERSACSIIIAFVVGALQYEDNVN	480
QY	481 PEMRPRVCMHYPRQGVSAKTVCGPVYCFTPSPVVVGTTDLGAPYTWGNETDVL	540
DB	481 PEMRPRVCMHYPRQGVSAKTVCGPVYCFTPSPVVVGTTDLGAPYTWGNETDVL	540
QY	541 LNSTRPPLGSMFCGTMMNNSGGYTKCGAPRCRTRADFNASTDLCPTDFCRKHPDTTYLK	600
DB	541 LNSTRPPLGSMFCGTMMNNSGGYTKCGAPRCRTRADFNASTDLCPTDFCRKHPDTTYLK	600
QY	601 CGSGPMLTPRCCLIDYPRRLMHPCTVANNYTIKIRMYVGVGVEHRLTAACNFTGRDCNLE	660
DB	601 CGSGPMLTPRCCLIDYPRRLMHPCTVANNYTIKIRMYVGVGVEHRLTAACNFTGRDCNLE	660
QY	661 RDSQSPLSLHSTTEWAIIPCSYSDLPALSTGLLHONIVDVQVNYGSLPALTKTYVM	720
DB	661 RDSQSPLSLHSTTEWAIIPCSYSDLPALSTGLLHONIVDVQVNYGSLPALTKTYVM	720
QY	721 EWTILLFLLLADARVCACTMMLILLGOAEALAEKLYILHAASASNGEFLYFVIFVAAW	780
DB	721 EWTILLFLLLADARVCACTMMLILLGOAEALAEKLYILHAASASNGEFLYFVIFVAAW	780
QY	781 YIKGRVVPATYSLTGLMSFSLILLALPOQAVAYDASVHGQIGALLVMTLTFTLTGXYK	840

DB	781 YIKGRVVPATYSLTGLMSFSLILLALPOQAVAYDASVHGQIGALLVMTLTFTLTGXYK	840
QY	841 TLLSRFLMWLCYLLTLGAMVOEWAAPMOWVGSNDGIIWAAVAIPVGVPEITTKMLAVL	900
DB	841 TLLSRFLMWLCYLLTLGAMVOEWAAPMOWVGSNDGIIWAAVAIPVGVPEITTKMLAVL	900
QY	901 GPAYLLKCALTRVPYFAVAHALLRNCTMARLHAGRYVQMALLAGRTGYIYDHLTPM	960
DB	901 GPAYLLKCALTRVPYFAVAHALLRNCTMARLHAGRYVQMALLAGRTGYIYDHLTPM	960
QY	961 SDMAASGLRDIAVAVEPIIFSPMEKVTIWMGAETACDIIHLGLPVSARLREVLGPAD	1020
DB	961 SDMAASGLRDIAVAVEPIIFSPMEKVTIWMGAETACDIIHLGLPVSARLREVLGPAD	1020
QY	1021 GYTSKWSLAPITAYAAOQTGILGTIVASMTGDKTEQAGEIOVLSTVTQSFLGTSISG	1080
DB	1021 GYTSKWSLAPITAYAAOQTGILGTIVASMTGDKTEQAGEIOVLSTVTQSFLGTSISG	1080
QY	1081 VLMTYVHAGNKTLAGSRGPTOMYSASGDLVGMPSPGTKSLPCTCGAVDLYLVRN	1140
DB	1081 VLMTYVHAGNKTLAGSRGPTOMYSASGDLVGMPSPGTKSLPCTCGAVDLYLVRN	1140
QY	1141 ADVIPARRGDKRGALLSPRLSTLKSSGSPVLCPRGAVGVRAAVCSRGVAKSIDFI	1200
DB	1141 ADVIPARRGDKRGALLSPRLSTLKSSGSPVLCPRGAVGVRAAVCSRGVAKSIDFI	1200
QY	1201 PVETLDIYTRSPPTSNDSTPRAVPOTVOGYLHAPTSGSKSTKVPVAAAGQYKVLVNP	1260
DB	1201 PVETLDIYTRSPPTSNDSTPRAVPOTVOGYLHAPTSGSKSTKVPVAAAGQYKVLVNP	1260
QY	1261 SVAATLFGAAYLSKAHGINPNI RTGVRVTYTGAPITVSYGKFLADGGCAGAYDIIICD	1320
DB	1261 SVAATLFGAAYLSKAHGINPNI RTGVRVTYTGAPITVSYGKFLADGGCAGAYDIIICD	1320
QY	1321 ECHAVDSTTLIGIGTVDLQAEATAGRLTVLATATPBGSVTTPHPIIEEVALGQSEIIPFY	1380
DB	1321 ECHAVDSTTLIGIGTVDLQAEATAGRLTVLATATPBGSVTTPHPIIEEVALGQSEIIPFY	1380
QY	1381 GRAIPLSYIKGGRHLIFCHSKKKCDELAALARGMLNVAAYRGLDVETITQGVVVVA	1440
DB	1381 GRAIPLSYIKGGRHLIFCHSKKKCDELAALARGMLNVAAYRGLDVETITQGVVVVA	1440
QY	1441 TDLMTGYTGDPFDSYIDCNVAVTQVVDPSLDPTEFTITQIYPODAVSRQRRGTGRRL	1500
DB	1441 TDLMTGYTGDPFDSYIDCNVAVTQVVDPSLDPTEFTITQIYPODAVSRQRRGTGRRL	1500
QY	1501 GIYRYVSTGERASGMFDSVILCECYDAGAAHYELTPSETVRLRAYVNTPGLPVCOQHLE	1560
DB	1501 GIYRYVSTGERASGMFDSVILCECYDAGAAHYELTPSETVRLRAYVNTPGLPVCOQHLE	1560
QY	1561 FWEAVFTGLTHIDAFILSOTKOSGENFAYLTAAYOATVCAKAKAPPSDVMKKCLTRLRP	1620
DB	1561 FWEAVFTGLTHIDAFILSOTKOSGENFAYLTAAYOATVCAKAKAPPSDVMKKCLTRLRP	1620
QY	1621 TLVGPPLLYRLGVTNENVTLTHPYTKYIATCMQADLEVMSTVWLAGVLAAYAYCLA	1680
DB	1621 TLVGPPLLYRLGVTNENVTLTHPYTKYIATCMQADLEVMSTVWLAGVLAAYAYCLA	1680
QY	1681 TGCVCITIGRLHINQRAVVAPEKEVLYEAFDEMEBCASPAALIIEBQRIAEMLKSKIQGL	1740
DB	1681 TGCVCITIGRLHINQRAVVAPEKEVLYEAFDEMEBCASPAALIIEBQRIAEMLKSKIQGL	1740
QY	1741 OQASKOADIDOPTVOASMPKVEQFAKMMNFISGIQTLAGLSTLPGNPAVASMMAFSA	1800
DB	1741 OQASKOADIDOPTVOASMPKVEQFAKMMNFISGIQTLAGLSTLPGNPAVASMMAFSA	1800
QY	1801 LTPSPSTTLILNLTIGMLASQIAPAGATGVVSGVGAAGSIGHKYLVDILLAGY	1860
DB	1801 LTPSPSTTLILNLTIGMLASQIAPAGATGVVSGVGAAGSIGHKYLVDILLAGY	1860
QY	1861 AGISGALVAFKIMSGEKSMEDEVNVLFGIISPGALVGVCAALLRHHVGBGAGVQWM	1920


```

Db 1861 AGISGALVAFKIMSGEKSEMEDVNNLPGILSPGALVWGVIICAILRRHVGBEGAVQNM 1920
Qy 1921 NRLIFASRGHNVAPTHVYTESDASORYTQLLGSLLTSLRLRIHNMWTEDECPICGGSM 1980
Db 1921 NRLIFASRGHNVAPTHVYTESDASORYTQLLGSLLTSLRLRIHNMWTEDECPICGGSM 1980
Qy 1981 LRDVDMWVTLLTDFKNMLTSLKLPMPGLPFVSCQKYGKGVAGTGMTTRCPGKANIS 2040
Db 1981 LRDVDMWVTLLTDFKNMLTSLKLPMPGLPFVSCQKYGKGVAGTGMTTRCPGKANIS 2040
Qy 2041 GNVRLGSMRITGPKTGMNIGTGFPIVNCCTEGOCVPPKAPNPKVAILWVAASBYAVTQH 2100
Db 2041 GNVRLGSMRITGPKTGMNIGTGFPIVNCCTEGOCVPPKAPNPKVAILWVAASBYAVTQH 2100
Qy 2101 GSYHITGLTDMNLKVPCLPSPPEFSWVDGQIHRFATPKPFPFDEVSFCVGLNSFV 2160
Db 2101 GSYHITGLTDMNLKVPCLPSPPEFSWVDGQIHRFATPKPFPFDEVSFCVGLNSFV 2160
Qy 2161 GSQLPCEPDPTDVLMSMLTDSHITAEFAARLARSGSPSEASASSASQLSAPSLRATCT 2220
Db 2161 GSQLPCEPDPTDVLMSMLTDSHITAEFAARLARSGSPSEASASSASQLSAPSLRATCT 2220
Qy 2221 THGKAYDMDVDANLPMGSDVTRISGSKVVLDSLDEWVERSDLEPSIPEEYMLPKKR 2280
Db 2221 THGKAYDMDVDANLPMGSDVTRISGSKVVLDSLDEWVERSDLEPSIPEEYMLPKKR 2280
Qy 2281 FPPALPAMARPDYNPLVESMWRPDYOPATVAGCALPPRRKPTTPPPRRRRVGLSEDSI 2340
Db 2281 FPPALPAMARPDYNPLVESMWRPDYOPATVAGCALPPRRKPTTPPPRRRRVGLSEDSI 2340
Qy 2341 GDAOLQAIKSPGPPSGDSGLSTGAGADSGSOTPPDELALSETGISISSMPLEGEIG 2400
Db 2341 GDAOLQAIKSPGPPSGDSGLSTGAGADSGSOTPPDELALSETGISISSMPLEGEIG 2400
Qy 2401 DEDLEPEQVPPPPQGGVAAFGSDSGSWSTSEEDDSVCCSWSYWGALITPPCSPE 2460
Db 2401 DEDLEPEQVPPPPQGGVAAFGSDSGSWSTSEEDDSVCCSWSYWGALITPPCSPE 2460
Qy 2461 EKLPIPNLSNLSLRHNNKYCTTTKSASIRAKKVFDRMQLVDSYDSDYKDIKLAASV 2520
Db 2461 EKLPIPNLSNLSLRHNNKYCTTTKSASIRAKKVFDRMQLVDSYDSDYKDIKLAASV 2520
Qy 2521 TARLLTMEBACOLTPPHSARSKYFGAKESVRLSGRAVNHISVWMDLLEDESEPIPTI 2580
Db 2521 TARLLTMEBACOLTTPHSARSKYFGAKESVRLSGRAVNHISVWMDLLEDESEPIPTI 2580
Qy 2581 MAKNEVFCVDPPTGGKKAARLLIVPDGLVRCERKALYDIQKLPQAVMGASVGFQYSPA 2640
Db 2581 MAKNEVFCVDPPTGGKKAARLLIVPDGLVRCERKALYDIQKLPQAVMGASVGFQYSPA 2640
Qy 2641 QRVRELLAKAMAKKDPMGSSYTRCPDSIVTERDITRTESTRAGSLPEEATTAHSLTE 2700
Db 2641 QRVRELLAKAMAKKDPMGSSYTRCPDSIVTERDITRTESTRAGSLPEEATTAHSLTE 2700
Qy 2701 RLYVGGPMFNSKGTGCGYRCRAGSVLTTSMGNITTCYKALAAKCAAGIAPMTLVCD 2760
Db 2701 RLYVGGPMFNSKGTGCGYRCRAGSVLTTSMGNITTCYKALAAKCAAGIAPMTLVCD 2760
Qy 2761 DLVVISSESGTEDEDERNLRAFTAMTRYSAAPGDPREPEYDELITSCSSNVVALGPG 2820
Db 2761 DLVVISSESGTEDEDERNLRAFTAMTRYSAAPGDPREPEYDELITSCSSNVVALGPG 2820
Qy 2821 RRRYVLTDPPTPIARAAMEYTRHSFVNSWNLGNIQYATTIARWVLMTHFSSILMAQT 2880
Db 2821 RRRYVLTDPPTPIARAAMEYTRHSFVNSWNLGNIQYATTIARWVLMTHFSSILMAQT 2880
Qy 2881 LDQNFEMFGAVYSPLDPAIIEIRLGLDAFSLHTYTPHELTRVASALRKLGAAPLR 2940
Db 2881 LDQNFEMFGAVYSPLDPAIIEIRLGLDAFSLHTYTPHELTRVASALRKLGAAPLR 2940
Qy 2941 AKSRRARAVRASLLISRGRAAVCGRYLFWMAVYTKLKLPLPEARLLDLSWFTVAGGG 3000
Db 2941 AKSRRARAVRASLLISRGRAAVCGRYLFWMAVYTKLKLPLPEARLLDLSWFTVAGGG 3000

```

```

Qy 3001 DIYHSRARRPRLILFGLLLFVGVGLFLTPAR 3033
Db 3001 DIYHSRARRPRLILFGLLLFVGVGLFLTPAR 3033

RESULT 2
POLG_HCVJ6
ID POLG_HCVJ6 STANDARD; PRT: 3033 AA.
AC P26650;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Genome polyprotein [contains: Capsid protein C (Core protein) (p22);
DE Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
DE (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (p21)
DE (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepatitisin)
DE (EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
DE NS4B (p27); Nonstructural protein NS5A (p56); Nonstructural protein
DE NS5B (p66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
OS Hepatitis C virus (isolate HC-J6) (HCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11113;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92044440; PubMed=1658196;
RA Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Linaka H., Machida A.,
RA Miyakawa Y., Mayumi M.;
RT "Nucleotide sequence of the genomic RNA of hepatitis C virus isolated
RT from a human carrier: comparison with reported isolates for conserved
RT and divergent regions."
RL J. Gen. Virol. 72:2697-2704 (1991).
CC -!- FUNCTION: The small proteins NS2A, NS2B, NS4A and NS4B are
CC hydrophobic, suggesting a possible membrane-related function. NS3
CC and NS5 may play a role in the viral RNA replication.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
CC precursor polyprotein, commonly with Asp or Glu in the P6
CC position. Cys or Thr in P1 and Ser or Ala in P1'.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -!- SUBUNIT: The virion of this virus is a nucleocapsid covered by a
CC lipoprotein envelope. The envelope consists of two proteins:
CC protein M and glycoprotein E. The nucleocapsid is a complex of
CC protein C and mRNA.
CC -!- SIMILARITY: Contains 1 peptidase S29 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D00944; BA00792.1; -
DR FIC: J01303; J01303.
DR HSP: P27958; IHB1.
DR MEROPS: S29.001; -.
DR MEROPS: U39.001; -.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR002522; HCV_capsid.
DR InterPro: IPR002521; HCV_core.
DR InterPro: IPR002519; HCV_env.
DR InterPro: IPR002531; HCV_NS1.
DR InterPro: IPR000745; HCV_NS4A.
DR InterPro: IPR001490; HCV_NS4B.
DR InterPro: IPR002668; HCV_NS5A.
DR InterPro: IPR002166; HCV_NS5B.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR009003; Pept_ser_Cys.
DR InterPro: IPR002518; Pept_U39_HCV_NS2.

```

```
DR InterPro: IPR004109; Peptidase_S29.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF01543; HCV_capsid; 1.
DR Pfam: PF01542; HCV_core; 1.
DR Pfam: PF01539; HCV_env; 1.
DR Pfam: PF01560; HCV_NS1; 1.
DR Pfam: PF01538; HCV_NS2; 1.
DR Pfam: PF02907; HCV_NS3; 1.
DR Pfam: PF01006; HCV_NS4a; 1.
DR Pfam: PF01001; HCV_NS4b; 1.
DR Pfam: PF01506; HCV_NS5a; 1.
DR Pfam: PF00271; Helicase_C; 1.
DR Pfam: PF00998; Viral_RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
KM ATP-binding; Coat protein; Envelope protein;
KM Glycoprotein; Helicase; Hydrolyase; Nonstructural protein; Polypotein;
KM RNA-directed RNA polymerase; Serine protease; Transferase;
KM Transmembrane.
FT INIT_MET 1 1
FT CHAIN 1 115 Removed from capsid protein C by the
FT CHAIN 116 191 cellular aminopeptidase.
FT CHAIN 192 383 Capsid protein C (Potential).
FT CHAIN 384 733 Matrix protein (Potential).
FT CHAIN 734 1010 Major envelope protein E (Potential).
FT CHAIN 1011 1619 Nonstructural protein NS1 (Potential).
FT CHAIN 1620 1866 Nonstructural protein NS2 (Potential).
FT CHAIN 1867 2017 Protease/helicase NS3 (Potential).
FT CHAIN 2018 3033 Nonstructural protein NS4a (Potential).
FT TRANSMEM 347 369 Nonstructural protein NS4b (Potential).
FT ACT_SITE 1087 1087 RNA-directed RNA polymerase (Potential).
FT ACT_SITE 1111 1111 Charge relay system (By similarity).
FT ACT_SITE 1169 1169 Charge relay system (By similarity).
FT NP_BIND 1234 1241 ATP (Potential).
FT SITE 1320 1323 DECH box.
FT CARBOHYD 196 196 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 209 209 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 234 234 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 305 305 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 417 417 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 423 423 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 430 430 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 448 448 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 477 477 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 534 534 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 542 542 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 558 558 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 578 578 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 627 627 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 649 649 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1091 1091 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 2038 2038 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 2811 2811 N-linked (GlcNAc...) (Potential).
SQ SSOURCE 3033 AA; 329165 MW; F957F5C1A273BE9E CRC64;

Query Match 98.2%; Score 15830; DB 1; Length 3033;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 2967; Conservative 30; Mismatches 36; Indels 0; Gaps 0;
```

```
Db 181 LLSCTTTPVSAAEVKNISTGTVMNTDCTNDSITWQLAAVLAHVGCVCERKGNASQCWI 240
QY PVSPPVAVORPGALTOGRTHIDIMVYMGATCSALYVVDLCGGWMLAAQMTIVSRQHM 300
Db 241 PVSPPVAVOQPGALTOGRTHIDIMVYMGATCSALYVVDLCGGWMLAAQMTIVSRQHM 300
QY VQDCNCSITPGTITGHRAMDMNMNNSPTATMILAYAMRVEVILDIISGHWGMEGLA 360
Db 301 VQDCNCSITPGTITGHRAMDMNMNNSPTATMILAYAMRVEVILDIISGHWGMEGLA 360
QY VQDCNCSITPGTITGHRAMDMNMNNSPTATMILAYAMRVEVILDIISGHWGMEGLA 360
Db 361 VQDCNCSITPGTITGHRAMDMNMNNSPTATMILAYAMRVEVILDIISGHWGMEGLA 360
QY YFSMGAAKXVIVILLILAAGVADARTHTVGSAAQOTGTLTFLPMGPRQKIQLVNTNSW 420
Db 361 YFSMGAAKXVIVILLILAAGVADARTHTVGSAAQOTGTLTFLPMGPRQKIQLVNTNSW 420
QY HINRFLALNCNDSLHTGFIASLFTYHSFNSGSCPERMSACRSIEAFRVGMALQYEDNTN 480
Db 421 HINRFLALNCNDSLHTGFIASLFTYHSFNSGSCPERMSACRSIEAFRVGMALQYEDNTN 480
QY HINRFLALNCNDSLHTGFIASLFTYHSFNSGSCPERMSACRSIEAFRVGMALQYEDNTN 480
Db 481 HINRFLALNCNDSLHTGFIASLFTYHSFNSGSCPERMSACRSIEAFRVGMALQYEDNTN 480
QY PEDMRPYCMHYPRQCGVSAKTYCGPVYCTPSPBVVGTDRIGAPTYTGENETDVEL 540
Db 481 PEDMRPYCMHYPRQCGVSAKTYCGPVYCTPSPBVVGTDRIGAPTYTGENETDVEL 540
QY LNSTRPPLGSMFGCTMNMNSGYTKTCGAPPCRTADPFASTDLLCPTDCFRKHDPPTYLK 600
Db 541 LNSTRPPLGSMFGCTMNMNSGYTKTCGAPPCRTADPFASTDLLCPTDCFRKHDPPTYLK 600
QY LNSTRPPLGSMFGCTMNMNSGYTKTCGAPPCRTADPFASTDLLCPTDCFRKHDPPTYLK 600
Db 541 LNSTRPPLGSMFGCTMNMNSGYTKTCGAPPCRTADPFASTDLLCPTDCFRKHDPPTYLK 600
QY CGSGWMLPRCLIDIPYLMHYPCVTNYTIKIRYVGVSHRLTLAACPFRGRCDNED 660
Db 601 CGSGWMLPRCLIDIPYLMHYPCVTNYTIKIRYVGVSHRLTLAACPFRGRCDNED 660
QY RDRSGLSPLSHSTTEWALIPCSYSDLPALSTGLLHQNIIVDVOFMVGLSPALTKYIVRW 720
Db 661 RDRSGLSPLSHSTTEWALIPCSYSDLPALSTGLLHQNIIVDVOFMVGLSPALTKYIVRW 720
QY EEWVILLFILLADARVACGLMILLIGQEALEKLVILHAASASCNGLFVYFVFAVAM 780
Db 721 EEWVILLFILLADARVACGLMILLIGQEALEKLVILHAASASCNGLFVYFVFAVAM 780
QY EEWVILLFILLADARVACGLMILLIGQEALEKLVILHAASASCNGLFVYFVFAVAM 780
Db 721 EEWVILLFILLADARVACGLMILLIGQEALEKLVILHAASASCNGLFVYFVFAVAM 780
QY YIKGIVPLATYSLTGLMSFSLILLALPOQAAVYADASVHGOIGALLVMTLFTTPGYK 840
Db 781 YIKGIVPLATYSLTGLMSFSLILLALPOQAAVYADASVHGOIGALLVMTLFTTPGYK 840
QY YIKGIVPLATYSLTGLMSFSLILLALPOQAAVYADASVHGOIGALLVMTLFTTPGYK 840
Db 781 YIKGIVPLATYSLTGLMSFSLILLALPOQAAVYADASVHGOIGALLVMTLFTTPGYK 840
QY TLRSFLMWLCYLLTLGSAVYQEMAPPMQVAGRGDIIMAVAFYPGVFDITKMLAVL 900
Db 841 TLRSFLMWLCYLLTLGSAVYQEMAPPMQVAGRGDIIMAVAFYPGVFDITKMLAVL 900
QY GPAYLLKQALTRVPYFVFAHALLRMCTMARHLAAGRYQMALALGRWTGYIYDHLTPM 960
Db 901 GPAYLLKQALTRVPYFVFAHALLRMCTMARHLAAGRYQMALALGRWTGYIYDHLTPM 960
QY GPAYLLKQALTRVPYFVFAHALLRMCTMARHLAAGRYQMALALGRWTGYIYDHLTPM 960
Db 901 GPAYLLKQALTRVPYFVFAHALLRMCTMARHLAAGRYQMALALGRWTGYIYDHLTPM 960
QY SDMAASGLRDIAVAVEPIIFSPMEKKVIVMGAEFAACDIIHGLPVSARLREVLGPD 1020
Db 961 SDMAASGLRDIAVAVEPIIFSPMEKKVIVMGAEFAACDIIHGLPVSARLREVLGPD 1020
QY SDMAASGLRDIAVAVEPIIFSPMEKKVIVMGAEFAACDIIHGLPVSARLREVLGPD 1020
Db 961 SDMAASGLRDIAVAVEPIIFSPMEKKVIVMGAEFAACDIIHGLPVSARLREVLGPD 1020
QY GYTSKQWMLAPITAYAOQTRGLLCTIVVSMGTGRDKTEQAGEIOVLSTVTQSLTGISIG 1080
Db 1021 GYTSKQWMLAPITAYAOQTRGLLCTIVVSMGTGRDKTEQAGEIOVLSTVTQSLTGISIG 1080
QY GYTSKQWMLAPITAYAOQTRGLLCTIVVSMGTGRDKTEQAGEIOVLSTVTQSLTGISIG 1080
Db 1021 GYTSKQWMLAPITAYAOQTRGLLCTIVVSMGTGRDKTEQAGEIOVLSTVTQSLTGISIG 1080
QY VLMTYTHAGANKTLAAGSGPYTQWSSAEGULVGMPSPTGKSLPECTCGAVDLVYLRN 1140
Db 1081 VLMTYTHAGANKTLAAGSGPYTQWSSAEGULVGMPSPTGKSLPECTCGAVDLVYLRN 1140
QY VLMTYTHAGANKTLAAGSGPYTQWSSAEGULVGMPSPTGKSLPECTCGAVDLVYLRN 1140
Db 1081 VLMTYTHAGANKTLAAGSGPYTQWSSAEGULVGMPSPTGKSLPECTCGAVDLVYLRN 1140
QY ADVIPARRRGDKRGALLSPRLSTLKSGSGGPPVLCPRHAAGVRAAACSGVAKSIDFI 1200
Db 1141 ADVIPARRRGDKRGALLSPRLSTLKSGSGGPPVLCPRHAAGVRAAACSGVAKSIDFI 1200
QY ADVIPARRRGDKRGALLSPRLSTLKSGSGGPPVLCPRHAAGVRAAACSGVAKSIDFI 1200
Db 1141 ADVIPARRRGDKRGALLSPRLSTLKSGSGGPPVLCPRHAAGVRAAACSGVAKSIDFI 1200
QY PVETLIDIVTRSPESDNTSPPAVPQTYGVYLAHPTGSGSKTKVPVAYAAQGVKVLVNP 1260
Db 1201 PVETLIDIVTRSPESDNTSPPAVPQTYGVYLAHPTGSGSKTKVPVAYAAQGVKVLVNP 1260
QY PVETLIDIVTRSPESDNTSPPAVPQTYGVYLAHPTGSGSKTKVPVAYAAQGVKVLVNP 1260
Db 1201 PVETLIDIVTRSPESDNTSPPAVPQTYGVYLAHPTGSGSKTKVPVAYAAQGVKVLVNP 1260
QY SYAATLGGAYLSKAGINENIRIGVRYVTYTGAPITTSYTGKFLADGGCAGADIIICD 1320
Db 1261 SYAATLGGAYLSKAGINENIRIGVRYVTYTGAPITTSYTGKFLADGGCAGADIIICD 1320
```

```

Db 1261 SVAATLGFAYLSKAGINPNIRGTGRTVTTGAPITVSTYGGKFLDAGCGAGADY111CD 1320
Qy 1321 ECHAVDSTTIIGTIVLDOAETAGVRLTVLAATPAGSVTPPHNIEVALGEGEIPFY 1380
Db 1321 ECHAVDSTTIIGTIVLDOAETAGVRLTVLAATPAGSVTPPHNIEVALGEGEIPFY 1380
Qy 1381 GRAIPLSYIKGGRHLIFCHSKKKCDELAALRGMLNSVAYYRGDVSIVPTQGVVVVA 1440
Db 1381 GRAIPLSYIKGGRHLIFCHSKKKCDELAALRGMLNSVAYYRGDVSIVPTQGVVVVA 1440
Qy 1441 TDALMTGTGDFDSYIDCNVAVTQVVDPSLDPFTTITQIYVQDAVSRSQRRGRGTGRGL 1500
Db 1441 TDALMTGTGDFDSYIDCNVAVTQVVDPSLDPFTTITQIYVQDAVSRSQRRGRGTGRGL 1500
Qy 1501 GIYRVVSIGERSGDFSVLCECYDAGAMVELPSETTVRLRAYFPTPGLPVQDHL 1560
Db 1501 GIYRVVSIGERSGDFSVLCECYDAGAMVELPSETTVRLRAYFPTPGLPVQDHL 1560
Qy 1561 FWEAVFTGLTHIDAFHLSQTKSGENFAVLTAQATVCARAKAPPSMDVMKCLTRLKP 1620
Db 1561 FWEAVFTGLTHIDAFHLSQTKSGENFAVLTAQATVCARAKAPPSMDVMKCLTRLKP 1620
Qy 1621 TLVGPFTLLYRLGVTNEVTLTHPVTKYIATCMQADLEVMSTWVLACGVLAAYCIA 1680
Db 1621 TLVGPFTLLYRLGVTNEVTLTHPVTKYIATCMQADLEVMSTWVLACGVLAAYCIA 1680
Qy 1681 TGCVCTIRLRHINQAVVAPDKEVLYEAPEDEBECASRALIEBQRIAEMLKSIQGL 1740
Db 1681 TGCVCTIRLRHINQAVVAPDKEVLYEAPEDEBECASRALIEBQRIAEMLKSIQGL 1740
Qy 1741 QOASQOADIPTVOASWPKEYQFPAKMMNFISGIQYLAGLSTLPGNPVAVSMAFSA 1800
Db 1741 QOASQOADIPTVOASWPKEYQFPAKMMNFISGIQYLAGLSTLPGNPVAVSMAFSA 1800
Qy 1801 LTPSLSTTILLNLILGMLASQIAPAGATGFVSGLVGAAGVSGIGLKVLDLILAGY 1860
Db 1801 LTPSLSTTILLNLILGMLASQIAPAGATGFVSGLVGAAGVSGIGLKVLDLILAGY 1860
Qy 1861 AGISALNAFKIMSEKSMEDVNNLPGILSPGALVYGVCAALIRRHVGBEGAVQOM 1920
Db 1861 AGISALNAFKIMSEKSMEDVNNLPGILSPGALVYGVCAALIRRHVGBEGAVQOM 1920
Qy 1921 NRLIFASRGHVAPTHVYTESDASQRTQLGSLTITSLRLNMWITTEDPFCGSGM 1980
Db 1921 NRLIFASRGHVAPTHVYTESDASQRTQLGSLTITSLRLNMWITTEDPFCGSGM 1980
Qy 1981 LRDVMDVCTLLTDFKMLTSLPKMPGLPFVSCQKGYKGVAGTGMTTRCPGAGNIS 2040
Db 1981 LRDVMDVCTLLTDFKMLTSLPKMPGLPFVSCQKGYKGVAGTGMTTRCPGAGNIS 2040
Qy 2041 GNVRLGSRITGPKTCMNIWQSTFPIVNCYTBEGQCPKPAVNPKVAILWRYASEVLETOH 2100
Db 2041 GNVRLGSRITGPKTCMNIWQSTFPIVNCYTBEGQCPKPAVNPKVAILWRYASEVLETOH 2100
Qy 2101 GSYHITGTGTDNLRKVPOLPSPERFSWVDGQIHRFAPTPKPRFDEVSFCVGLNSFV 2160
Db 2101 GSYHITGTGTDNLRKVPOLPSPERFSWVDGQIHRFAPTPKPRFDEVSFCVGLNSFV 2160
Qy 2161 GSQLPCEDEPDTVLMSMLTDBSHITAAARLARLARGSPBSASSASQSLAPSLRATCT 2220
Db 2161 GSQLPCEDEPDTVLMSMLTDBSHITAAARLARLARGSPBSASSASQSLAPSLRATCT 2220
Qy 2221 THGKAYVDVMDVANIPLMGDVTIRISSGSVNVVLDLDPVVERSDLEPSIPSEYMLPKKR 2280
Db 2221 THGKAYVDVMDVANIPLMGDVTIRISSGSVNVVLDLDPVVERSDLEPSIPSEYMLPKKR 2280
Qy 2281 FPPALPAPAPDYNPLVESWTKRPPYQATVAGCALPPRKPTPTPPRRRRRTVGLSEDSI 2340
Db 2281 FPPALPAPAPDYNPLVESWTKRPPYQATVAGCALPPRKPTPTPPRRRRRTVGLSEDSI 2340
Qy 2341 GDALQOLAIKSFQPPSGDSGLSTGAGAADSGQTPPDELALSETGSISSMPLEGEIG 2400
Db 2341 ADALQOLAIKSFQPPSGDSGLSTGAGAADSGSRTPPDELALSETGSISSMPLEGEIG 2400

```

```

Qy 2401 DPLEPEQVEPQPPQGGVAAAGSDSGSWSTCSBEDSDVCCSNYSWTGALLTPCSPEE 2460
Db 2401 DPLEPEQVEPQPPQGGVAAAGSDSGSWSTCSBEDSDVCCSNYSWTGALLTPCSPEE 2460
Qy 2461 EKLPIPLNSLRLRYHNKVCYCTTKSASLRAKKTVPEDMOYLDYSYDVKDIKLAASKV 2520
Db 2461 EKLPIPLNSLRLRYHNKVCYCTTKSASLRAKKTVPEDMOYLDYSYDVKDIKLAASKV 2520
Qy 2521 TAPLITTEACQLPHPHARSKYGGAKEVNSLSGRAVNHISKWKDLLEPSEPTPTTI 2580
Db 2521 TAPLITTEACQLPHPHARSKYGGAKEVNSLSGRAVNHISKWKDLLEPSEPTPTTI 2580
Qy 2581 MAKNEVFCVDPKGGKKAARLIYVPDIGVRCEKMAVDITOKLPOAVMGASVGFQYSPA 2640
Db 2581 MAKNEVFCVDPKGGKKAARLIYVPDIGVRCEKMAVDITOKLPOAVMGASVGFQYSPA 2640
Qy 2641 QVEBELKAMAEKKDPMGFSYDTRCFDSTVTERDIRTEESIVRACSLPEAHTAIHSLTE 2700
Db 2641 QVEBELKAMAEKKDPMGFSYDTRCFDSTVTERDIRTEESIVRACSLPEAHTAIHSLTE 2700
Qy 2701 RLHYGGEFNSKQGTGCRRCRAGSVLTTSMGNTTTCYVKALAAKAGIILAPTMLVCGD 2760
Db 2701 RLHYGGEFNSKQGTGCRRCRAGSVLTTSMGNTTTCYVKALAAKAGIILAPTMLVCGD 2760
Qy 2761 DLVVISSEGTBEDEBNLRAPTEAMTRYSAPGDPPEPYDLELITSSSNVSVALGFG 2820
Db 2761 DLVVISSEGTBEDEBNLRAPTEAMTRYSAPGDPPEPYDLELITSSSNVSVALGFG 2820
Qy 2821 RRRYVLTEDPTPIARAAMETVRHSVNSWLGNIIOVAPTIWARMVMTFFESILMAODT 2880
Db 2821 RRRYVLTEDPTPIARAAMETVRHSVNSWLGNIIOVAPTIWARMVMTFFESILMAODT 2880
Qy 2881 LDONINFEWGAVSVSPLDPAIIEHAGDAPSLHTYTHELTRVASALPKGAPPLR 2940
Db 2881 LDONINFEWGAVSVSPLDPAIIEHAGDAPSLHTYTHELTRVASALPKGAPPLR 2940
Qy 2941 AKMSRARAVRASLSIRGGRAAVCGRYLPNMAVKTGLTLPPEARLLDLSWFTVAGAGG 3000
Db 2941 AKMSRARAVRASLSIRGGRAAVCGRYLPNMAVKTGLTLPPEARLLDLSWFTVAGAGG 3000
Qy 3001 DIYHSVRARPRLILFGILLFVGVGLFLPAR 3033
Db 3001 DIYHSVRARPRLILFGILLFVGVGLFLPAR 3033

RESULT 3
ID 0909A7 PRELIMINARY; PRT: 3033 AA.
AC 0909A7;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepcivirns.
OX NCBI_TaxID=11103;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=21361470; PubMed=11468731; DOI=10.1002/jmv.1073;
RA Kurihara C., Ishiyama N., Nishiyama Y., Fukushima S., Kageyama T.,
RA Kageyama K., Miura S.;
RT Molecular characterization of hepatitis C virus genotype 2a from the
RT entire sequences of four isolates.;
RT J. Med. Virol. 64:466-475(2001).
DR EMBL; AF159005; AAF25613.1; -.
DR HSSP; Q8JYS1; ICWX.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.

```

DR GO:00031723; F:RNA binding; IEA.
DR GO:00033966; F:RNA-directed RNA polymerase activity; IEA.
DR GO:00082236; F:serine-type peptidase activity; IEA.
DR GO:00051598; F:structural molecule activity; IEA.
DR GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO:0006350; P:transcription; IEA.
DR GO:0019079; P:viral genome replication; IEA.
DR GO:0019087; P:viral transformation; IEA.
DR InterPro: IPR000345; Cyclic_heme_BS.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR011545; DEAD/DEAH_N.
DR InterPro: IPR002522; HCV capsid.
DR InterPro: IPR002521; HCV_core.
DR InterPro: IPR002519; HCV_env.
DR InterPro: IPR002531; HCV_NS1.
DR InterPro: IPR000745; HCV_NS4A.
DR InterPro: IPR001490; HCV_NS4B.
DR InterPro: IPR002868; HCV_NS5A.
DR InterPro: IPR002166; HCV_RdRp.
DR InterPro: IPR004109; Peptidase_S29.
DR InterPro: IPR009003; Pept_Ser_Cys.
DR InterPro: IPR002518; Pept_039_HCV_NS2.
DR InterPro: IPR007095; RNA pol_DS_B.
DR InterPro: IPR007094; RNA pol_DsViral.
DR Pfam: PF01543; HCV_capsid; 1.
DR Pfam: PF01542; HCV_core; 1.
DR Pfam: PF01539; HCV_env; 1.
DR Pfam: PF01560; HCV_NS1; 1.
DR Pfam: PF01538; HCV_NS2; 1.
DR Pfam: PF02907; HCV_NS3; 1.
DR Pfam: PF01006; HCV_NS4a; 1.
DR Pfam: PF01001; HCV_NS4b; 1.
DR Pfam: PF01506; HCV_NS5a; 1.
DR Pfam: PF00998; Viral_RdRp; 1.
DR SMART: SMO0487; DEXDC; 1.
DR PROSITE: PS00190; CYTOCHROME C; UNKNOWN.1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein; Polyprotein; Transmembrane.
KW POLYPEPTIDE 3033 AA; 329222 MW; 21492386CA0D5D8C CRC64;

Query Match	95.7%	Score 15426	DB 2	Length 3033
Best Local Similarity	94.8%	Pred. No. 0		
Matches 2874	Conservative 72	Mismatches 87	Indels 0	Gaps 0
QY	1	MSINPKPORKTKRNTNRPRPODVKEFGGGQIVGGVYLLPRRGRLGVARTRTKTSERSQGRG	60	
Db	1	MSINPKPORKTKRNTSRPRPODVKEFGGGQIVGGVYLLPRRGRLGVARTRTKTSERSQGRG	60	
QY	61	RROPTRPKRBSRTGSKMGKPGYWPPIYXNEGGLGMAWLLSPRGSRSPWGPNDPRHRSRVG	120	
Db	61	RROPTRPKRBSRTGSKMGKPGYWPPIYXNEGGLGMAWLLSPRGSRSPWGPNDPRHRSRVG	120	
QY	121	KVIDTLTGCFADLMGVIYVCGAGLGGVBARALAHGRVLEDEGVAFATGMIPGCSFSIFLLA	180	
Db	121	KVIDTLTGCFADLMGVIYVCGAGLGGVBARALAHGRVLEDEGVAFATGMIPGCSFSIFLLA	180	
QY	181	LLSCTTTPVSAAEVKNISTGYMTNDCTNDSITWLOQAIVLHPGCPVCEKXVNAQSCWI	240	
Db	181	LLSCTTTPVSAAYQVKNISDSYMTNDCTNDSITWLOQAIVLHPGCPVCEKXGKNTSRQCI	240	
QY	241	PVSPVPAVQVRGALVQGLRTHLDVYVMSATLCSALYYGDLGGWVLAQMFIVSPOHMF	300	
Db	241	PVSPVPAVQVRGALVQGLRAHIDVYVMSATLCSALYYGDLGGWVLAQMFIVSPOHMF	300	
QY	301	VODNCSTIYPGTTTGHRAAMDMMNMSPATWILAYAMRVBEVIDIISGAHGYMFGIA	360	
Db	301	VOECNCSTIYPGATITGHRAAMDMMNMSPATWILAYAMRVBEVIDIISGAHGYMFGIA	360	
QY	361	YFSNQGAAKXVVYVLLLAAGVADARHTYVGGSAQOTGGLTSLFMDGPRQKTQVNTNCSW	420	
Db	361	YFSNQGAAKXVVYVLLLAAGVADARHTSITAGSAHAHTSGIAGLFTISGARQNTQVINTNSW	420	
QY	421	HINRPAIACNDSLHTGFTASLFTYTSFMSQSPCEPMSACRSLEAFRVGMALQYEDNTN	480	

Dd	421	HINRTALCNDLSJLTGFIASLPIYTYRENSSCPEBLSACRGIQAFRIQMLGRLREDNTN	480
Qy	481	PEDMRPYCMWHPRQCGVVSAAKTCGPPYCTPSPVVWGTDRIGAPTYTGENETDVL	540
Dd	481	PEDMRPYCMWHPRQCGVVSAAKTCGPPYCTPSPVVWGTDRIGAPTYTGENETDVI	540
Qy	541	LNSRTPPLGSMFGCTMWNSSGYTTCGAPPCRTADPFASITDLCPDPCFKHPDPTYLK	600
Dd	541	LNSRTPPGGSMFGCTMWNSTGFTXTCGAPPCRIADPFASNDLLCPDPCFKHPDPTYLK	600
Qy	601	CGSGWMLPRLCIYPRLYMHYPCGVVNTIKIRMYGVGVHRLTAACNFRGDRCNLED	660
Dd	601	CGSGWMLPRLCIYPRLYMHYPCCTINTIKIRMYGVGVHRLTAACNFRGDRCNLED	660
Qy	661	RDRSOLSELHSTTEWAILPCSYSDLPALSTGLHLHONIDVOFMYGLSPALTKYIRW	720
Dd	661	RDRSOLSELHSTTEWAILPCSYSDLPALSTGLHLHONIDVOFMYGLSPALTKYIRW	720
Qy	721	EWVILLFLLADARVACLMWILLGOEALEBKVLIIHAASAAACNGFLYFVFFVAW	780
Dd	721	EWVILLFLLADARVACLMWILLGOEALEBKVLIIHAASAAACNGFLYFVFFVAW	780
Qy	781	YIKGRVPLATYSITGLMSFSLLLALPQOXYAADAASHGOIGALLVMITLFTLPBXY	840
Dd	781	YIKGRAVPLAYSITGLMPFCLLLALPQOXYAADAASHGOIGALLVMITLFTLPBXY	840
Qy	841	TLSRFLMWLCYLLTLGEMAOEAMPQVAGRGDIWAVALIFPGVVPDITKMLLVL	900
Dd	841	TLSRSLMWLCYLLTLGEMAOEAMPQVAGRGDIWAATMCPGVVPDITKMLLVL	900
Qy	901	GPAVLLKCALTRVYPVRAHALLRMCWMAHLAGRVYOMALLAGRWGTGYIYDHLTPM	960
Dd	901	GPAVLLRDLATRVYPVRAHALLRLCTVWRHLAGRVYOMALLAGRWGTGYIYDHLSPM	960
Qy	961	SDMAASGRLDAVAVEPIIFSPMEKRYIWMAGETPAACDILHGLPYVARLIREVLLGPAD	1020
Dd	961	SDMAASGRLDAVAVEPIIFSPMEKRYIWMAGETPAACDVLHGLPVSARLIREVLLGPAD	1020
Qy	1021	GYSKGMSGLAPIRAYAOQTRGLLGGTYIVSMTRGDKTEQAGEIYVLSVTQSPFGTSSISG	1080
Dd	1021	GYSKGMRLAPIRAYAOQTRGLLGLAIVSMTRGDKTEQAGEIYVLSVTQSPFGTSSISG	1080
Qy	1081	VLMTVYHAGNKTLAGSRPVTOMYSSAEGDLVGMPSBPSTKSLDEPTCGAVDLYLVYRN	1140
Dd	1081	VLMTVYHAGNKTLAGSRPVTOMYSSAEGDLVGMPSBPSTKSLDEPTCGAVDLYLVYRN	1140
Qy	1141	ADVIPARRRGKRGALLSPRLSTLKSSGGPVLCPRGHANVGPRAAVCSGVAKSIDFI	1200
Dd	1141	ADVIPARRRGKRGALLSPRLSTLKSSGGPVLCPRGHAGIIPRAACSRSGVAKSIDFI	1200
Qy	1201	PVELTLDIVTRSPTESDNSTPRAVQOTYOVGLHAPTGSGSKTKPVVAAGAOGYKVLVNP	1260
Dd	1201	PVEALDVTRSPTESDNSTPRAVPOTYOVGLHAPTGSGSKTKPVVAAGAOGYKVLVNP	1260
Qy	1261	SVAATLIGCALISRAHGINPNIIRTVGVTVTTGAPITYSTYKFLADGGCAGAGAYDIICD	1320
Dd	1261	SVAATLIGGALISRAHGINPNIIRTVGVTVTTGEBISTYTYKFLADGGCAGAGAYDIICD	1320
Qy	1321	ECHAVDSTTIIIGIGVLDQAEAGVRLTVLATATPPGCVTTPHPNIEEVALGOEGEIPFY	1380
Dd	1321	ECHAVDATTTIIGVGVLDQAEAGVRLTVLATATPPGCVTTPHPNIEEVALGOEGEIPFY	1380
Qy	1381	GRAIPPLSTIKGRHLIFCHSKKKCDELAALRGMGINSVAYYRGLDVSVITPQGVVVVA	1440
Dd	1381	GRAIPPLSTIKGRHIFCHSKKKCDELSAALRSGLNANVAYYRGLDVSVITPQGVVVVA	1440
Qy	1441	TDALMTGTGPDSVIDCNVAVTVQVDPSLDPTFTITTOIYPOPAVSRSGRRGRGRRL	1500
Dd	1441	TDALMTGTGPDSVIDCNVAVTVQVDPSLDPTFTITTOIYPOPAVSRSGRRGRGRRL	1500
Qy	1501	GIYRVVSTGERASGMFDSVVLCECYDAGAAWELTPSETTVRLRAYFTPGI.PVCOHLE	1560

Db 1501 G1YRVSTGERASGMFDSVILCECYDAGAAWYELTPSETTVRLRAYFNTPGLPVCQDHL 1560
QY 1561 FMEAVFTGLTHDAHFLSOTKSGENFAVLTAYATVCARAKAPPSMDVMKCLTRLRKP 1620
Db 1561 FMEAVFTGLTHDAHFLSOTKSGENFAVLTAYATVCARAKAPPSMDVMKCLTRLRKP 1620
QY 1621 TLVGFPTLLYRLGVTNEVTLTHPATKYIATCMQADLEVMSTWVLAAVAAVCL 1680
Db 1621 TLVGFPTLLYRLGVTNEVTLTHPATKYIATCMQADLEVMSTWVLAAVAAVCL 1680
QY 1661 TGCVCITGRHLINORAVAPDXEVLVEAFDEMEBCASRAALIIEGQRIAEMLKSKIQGL 1740
Db 1661 TGCVCITGRHLINORAVAPDXEVLVEAFDEMEBCASRAALIIEGQRIAEMLKSKIQGL 1740
QY 1741 QOASQOADDIOPTVOASMPKXEQFMAKHMNFISGIQYIAGISTPGNPVAVASMAFSA 1800
Db 1741 QOASQOADDIOPTVOASMPKXEQFMAKHMNFISGIQYIAGISTPGNPVAVASMAFSA 1800
QY 1801 LTPSTSTTLLNLTILGMLASQIAPAGATGFVVSGLVGAIVGSIGLKVLVDILAGX 1860
Db 1801 LTPSTSTTLLNLTILGMLASQIAPAGATGFVVSGLVGAIVGSIGLKVLVDILAGX 1860
QY 1861 AGISGALVAFKIMSGEKSPMEDVNLPGIISPGALVGVICAILRRHVGBEGAVQM 1920
Db 1861 AGISGALVAFKIMSGEKSPMEDVNLPGIISPGALVGVICAILRRHVGBEGAVQM 1920
QY 1921 NRLIFARGNHVAATHVATESDASQRTQLIGSITSLRLNMTTEDPITCGGW 1980
Db 1921 NRLIFARGNHVAATHVATESDASQRTQLIGSITSLRLNMTTEDPITCGGW 1980
QY 1981 LRDVADWCTTILDFKNMLTSLPKMPGLPFVSCOKGKGVAGTGTITTCPCGANIS 2040
Db 1981 LRDVADWCTTILDFKNMLTSLPKMPGLPFVSCOKGKGVAGTGTITTCPCGANIS 2040
QY 2041 GNVRLGSKRITGPKTCMNIWQSTPPINCYTEGQCVKPAKPAKVAIWRVAASEVAVTGH 2100
Db 2041 GNVRLGSKRITGPKTCMNIWQSTPPINCYTEGQCVKPAKPAKVAIWRVAASEVAVTGH 2100
QY 2101 GSYHITGLTMDNLKVPQCLSPERFSWVDGQIHRFAPTRPFREDVSCVGLNSPV 2160
Db 2101 GSYHITGLTMDNLKVPQCLSPERFSWVDGQIHRFAPTRPFREDVSCVGLNSPV 2160
QY 2161 GSQLECPDEPPTDVMMSMLTDPSHITAETAARLARLARGSPSSASSASQSLAPSLRATC 2220
Db 2161 GSQLECPDEPPTDVMMSMLTDPSHITAETAARLARLARGSPSSASSASQSLAPSLRATC 2220
QY 2221 THGKAYDVMDVANDLPMGADVTRIESGKVVLDSLDPWEERSDLPSIPSEYMLPKKR 2280
Db 2221 THGKAYDVMDVANDLPMGADVTRIESGKVVLDSLDPWEERSDLPSIPSEYMLPKKR 2280
QY 2281 PPPALPAAARPDYNPLVESMKRPDYQATVAGCALPPPKRTPTPPRRRRRVGISEDSI 2340
Db 2281 PPPALPAAARPDYNPLVESMKRPDYQATVAGCALPPPKRTPTPPRRRRRVGISEDSI 2340
QY 2341 GDALQOLAIKAFGQPLSGDSGLSTGADABDSGSTRTPEDSALSETGSISSMPLEGEG 2400
Db 2341 GDALQOLAIKAFGQPLSGDSGLSTGADABDSGSTRTPEDSALSETGSISSMPLEGEG 2400
QY 2401 DEDLEPEQVEPQPPQGGVAAAGSDSGSWTCSSEBDSVVCSSMSYWTGALITPCSPBE 2460
Db 2401 DEDLEPEQVEPQPPQGGVAAAGSDSGSWTCSSEBDSVVCSSMSYWTGALITPCSPBE 2460
QY 2461 EKLPIPNPISNSILRRHNKYCTTTSASISARAKVTFDRQVUTDYYDSVLKIKLAASV 2520
Db 2461 EKLPIPNPISNSILRRHNKYCTTTSASISARAKVTFDRQVUTDYYDSVLKIKLAASV 2520
QY 2521 TARTLTMEBAQOLTPPHSARSKYGAKEVRLSGRAVNHISVWKDLLEDSETPIPTTI 2580
Db 2521 TARTLTMEBAQOLTPPHSARSKYGAKEVRLSGRAVNHISVWKDLLEDSETPIPTTI 2580
QY 2581 MAKNEVFCVDPPTGGKKAARLIVYDPLGVRCERKALYDIQKLPQAVMGASGYFOYSPA 2640
Db 2581 MAKNEVFCVDPPTGGKKAARLIVYDPLGVRCERKALYDIQKLPQAVMGASGYFOYSPA 2640

QY 2641 ORVEFLKAMAEKKDPMGFSYDTRCFDSTVTERDITRTEESTIYRACSLPEBAHTAHSLTE 2700
Db 2641 ORVEFLKAMAEKKDPMGFSYDTRCFDSTVTERDITRTEESTIYRACSLPEBAHTAHSLTE 2700
QY 2701 RUVYGMFENSQKQTCGRRCRASGVLTSMGNTITCYVAKLAACKAAGIAPTMVLCGD 2760
Db 2701 RUVYGMFENSQKQTCGRRCRASGVLTSMGNTITCYVAKLAACKAAGIAPTMVLCGD 2760
QY 2761 DLVISESQGTIEEDERNIRAFTEAMTRYSAPPGDPRREYDELEITSCSSNVVALGPG 2820
Db 2761 DLVISESQGTIEEDERNIRAFTEAMTRYSAPPGDPRREYDELEITSCSSNVVALGPG 2820
QY 2821 RRRYLTDPPTPIRAAMETVRHSFVNSWGNITQYPTIWARAVLMTHTFSSILMAODT 2880
Db 2821 RRRYLTDPPTPIRAAMETVRHSFVNSWGNITQYPTIWARAVLMTHTFSSILMAODT 2880
QY 2881 LDONLNFEMYGAVSVSPLDLPATIERLHGLDAFSLHTYTPHETLRVASALRKLGAAPLR 2940
Db 2881 LDONLNFEMYGAVSVSPLDLPATIERLHGLDAFSLHTYTPHETLRVASALRKLGAAPLR 2940
QY 2941 AMKSPARAVRASLSIRGGRAVCGRYLRNMAVTKLKTPLPEARLLDLSWFTVGAAGG 3000
Db 2941 AMKSPARAVRASLSIRGGRAVCGRYLRNMAVTKLKTPLPEARLLDLSWFTVGAAGG 3000
QY 3001 DIYHSVRARPRLILFGLLLFVGVGLFLPLPAR 3033
Db 3001 DIYHSVRARPRLILFGLLLFVGVGLFLPLPAR 3033

RESULT 4
0991B6 PRELIMINARY; PRT: 3033 AA.
AC 0991B6;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N. A.
RX MEDLINE=23316767; PubMed=11424123; DOI=10.1002/jmv.1055;
RA Kato T., Furusaka A., Miyamoto M., Date T., Yasui K., Hiramoto J.,
RA Nagayama K., Tanaka T., Makita T.;
RT "Sequence analysis of hepatitis C virus isolated from a fulminant
RT hepatitis patient."
RL J. Med. Virol. 64:334-339(2001).
DR EMBL: AB047641; BAB32874.1; -.
DR HSSP; Q8JYS1; 1CWK.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral genome transformation; IEA.
DR InterPro; IPR000345; CyC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.

DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRp.
DR InterPro; IPR004109; Peptidease_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U99_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSV1r.
DR Pfam; PF01543; HCV_core1; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; Viral_RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3033 AA; 328780 MW; 9AB16077CF46517B CRC64;

Query Match 95.5%; Score 15394; DB 2; Length 3033;
Best Local Similarity 94.8%; Pred. No. 0;
Matches 2875; Conservative 74; Mismatches 84; Indels 0; Gaps 0;

QY 1 MSTNPQRKTKRNTRRPODYKEPGGGQIVGGVYLLPRGRPLGVRATRKTSERSQPRG 60
DB 1 MSTTPRQKTKRNTRRPODYKEPGGGQIVGGVYLLPRGRPLGVRATRKTSERSQPRG 60
QY 61 RRQPIKDRRSSTGSKGKPGYPMWLYGNELGWAGWLLSPRSGRPSGPNDRHRSRNVG 120
DB 61 RRQPIKDRRSSTGSKGKPGYPMWLYGNELGWAGWLLSPRSGRPSGPNDRHRSRNVG 120
QY 121 KYIDTITCGFADLMGIIPIVVGAPLGVAALAHGVLEDDGVNFATGNI.PGCSFSIFLLA 180
DB 121 KYIDTITCGFADLMGIIPIVVGAPLGVAALAHGVLEDDGVNFATGNI.PGCSFSIFLLA 180
QY 181 LLSCTITPVSAAEVKNIISTGYWVTNDCTNDSITWQLOAAVLHPGCVPECKGNASQCM 240
DB 181 LLSCTITPVSAAEVKNIISTGYWVTNDCTNDSITWQLOAAVLHPGCVPECKGNASQCM 240
QY 241 PVS PNVAVORPGALITGELRTHIDMVMSATLCSALVYGDLCGGVLAQMFIYSPOHMF 300
DB 241 PVS PNVAVORPGALITGELRTHIDMVMSATLCSALVYGDLCGGVLAQMFIYSPOHMF 300
QY 301 VDDCNCSTIPGTTTGRAMDMMMNNSPTATMILAYMRVPEYIDITISGAHWGVFGIA 360
DB 301 VDDCNCSTIPGTTTGRAMDMMMNNSPTATMILAYMRVPEYIDITISGAHWGVFGIA 360
QY 361 YESMOGAMKVVYITLLAAGVDPARTHTVGGSAAGCTGRLTSLFDMPGRKIDLVNTNGS 420
DB 361 YESMOGAMKVVYITLLAAGVDPARTHTVGGSAAGCTGRLTSLFDMPGRKIDLVNTNGS 420
QY 421 HINRTALNCNDSLHTSFIALFYTHSPNSGCEERNASACRSIEAFRVGALQYEDNVN 480
DB 421 HINRTALNCNDSLHTSFIALFYTHSPNSGCEERNASACRSIEAFRVGALQYEDNVN 480
QY 481 PEMRMRYCHYPRROGCVSAKTVCGPVYCTFPSPVVVGTTDRGLAPYTTWGENETDVL 540
DB 481 PEMRMRYCHYPRROGCVSAKTVCGPVYCTFPSPVVVGTTDRGLAPYTTWGENETDVL 540
QY 541 LNSTRPGLGSWFECTMNNSSGYTKTGAPRCRTRADFNASTDILCTDCEFRKHPTTYLK 600
DB 541 LNSTRPGLGSWFECTMNNSSGYTKTGAPRCRTRADFNASTDILCTDCEFRKHPTTYLK 600
QY 601 CGSGPWLTRCLIDYPRLMHYPCVTNNTYIFKIRMTVGVGVEHRLTAACNFTGDRCNLED 660
DB 601 CGSGPWLTRCLIDYPRLMHYPCVTNNTYIFKIRMTVGVGVEHRLTAACNFTGDRCNLED 660
QY 661 RDRSQSLPLSHSTTEWALLPCTYSIDLPAISTGLLHLHONIVDVQVMYGLSPALTKYIWR 720

DB 661 RDRSQSLPLSHSTTEWALLPCTYSIDLPAISTGLLHLHONIVDVQVMYGLSPALTKYIWR 720
QY 721 EMVILTFLLADARCACTMMLIILGQAEALBETKVLIIHAASASCNELFVFIPEVAVM 780
DB 721 EMVILTFLLADARCACTMMLIILGQAEALBETKVLIIHAASASCNELFVFIPEVAVM 780
QY 781 YIKGRVPLATYSLTGLWSFSLLLALPQAAVAVDASVHGOIGALLVMITLFTLTPGYK 840
DB 781 YIKGRVPLATYSLTGLWSFSLLLALPQAAVAVDASVHGOIGALLVMITLFTLTPGYK 840
QY 841 TLISRFLLWLCYLLTLGSAHQEWAAPMVQVGRDGIIMAVAIPEYGVFPIITKMLAVL 900
DB 841 TLISRFLLWLCYLLTLGSAHQEWAAPMVQVGRDGIIMAVAIPEYGVFPIITKMLAVL 900
QY 901 GPAYILKCALTRVPFVFAHALLRCTMARHLAAGRYQOMALLAGRTGYIYDHLTPM 960
DB 901 GPAYILKCALTRVPFVFAHALLRCTMARHLAAGRYQOMALLAGRTGYIYDHLTPM 960
QY 961 SDMAASGLRDLAVAVEPIIFSPMEKRVIVMGAETACGDILHGLPVSARLGREVLLGPAD 1020
DB 961 SDMAASGLRDLAVAVEPIIFSPMEKRVIVMGAETACGDILHGLPVSARLGREVLLGPAD 1020
QY 1021 GYTSKGMGLAPITAYAQOTRGLGCTIYVSMTRGRDKTEQAGEIOVLSTVTSFLGTSISG 1080
DB 1021 GYTSKGMGLAPITAYAQOTRGLGCTIYVSMTRGRDKTEQAGEIOVLSTVTSFLGTSISG 1080
QY 1081 VLMATVYHAGNKTLAGSGRPTOMYSSAEGDLVGPSPGPKLSLEPCTCGAVDLYLVTRN 1140
DB 1081 VLMATVYHAGNKTLAGSGRPTOMYSSAEGDLVGPSPGPKLSLEPCTCGAVDLYLVTRN 1140
QY 1141 ADVIPARRRGRKGLALSPRLSTLKGSSGGPVLCPRHAGVFRPAVCSRGVASKIDFI 1200
DB 1141 ADVIPARRRGRKGLALSPRLSTLKGSSGGPVLCPRHAGVFRPAVCSRGVASKIDFI 1200
QY 1201 PYETIDITRRSPPTSNDNSTPPAVPQTYGVGLHAPVTSGSKSTKVPVAAAGYKVLVNP 1260
DB 1201 PYETIDITRRSPPTSNDNSTPPAVPQTYGVGLHAPVTSGSKSTKVPVAAAGYKVLVNP 1260
QY 1261 SVAATLFGAYLSKXHGINPNIIRGVRVTVTGAPITVSTYKPLADGGCAGAYDIIICD 1320
DB 1261 SVAATLFGAYLSKXHGINPNIIRGVRVTVTGAPITVSTYKPLADGGCAGAYDIIICD 1320
QY 1321 ECHAVDSTIIIGIVLDOAETAGVRLTVLATAATPPGSVTTPHNIEBVALGOEGELIPY 1380
DB 1321 ECHAVDSTIIIGIVLDOAETAGVRLTVLATAATPPGSVTTPHNIEBVALGOEGELIPY 1380
QY 1381 GRAIPLSTIKGGRHLIFCHSKKKCDLAAALRGMLNSVAYYRGIDVSVITPOGDVVVVA 1440
DB 1381 GRAIPLSTIKGGRHLIFCHSKKKCDLAAALRGMLNSVAYYRGIDVSVITPOGDVVVVA 1440
QY 1441 TDALMTGYTGDPSVIDCNVAVTQVVDPSLPFTITTOIPDPAVSHSQRGRTGRRL 1500
DB 1441 TDALMTGYTGDPSVIDCNVAVTQVVDPSLPFTITTOIPDPAVSHSQRGRTGRRL 1500
QY 1501 GIYRVYSTGERASGMFDSVVLCECYDAGAAMVELTPSETTVRLRAYFNTPLPYCQDHL 1560
DB 1501 GIYRVYSTGERASGMFDSVVLCECYDAGAAMVELTPSETTVRLRAYFNTPLPYCQDHL 1560
QY 1561 FWEAVFTGLTHIDAFISQTOAGDNFAYLTAQATVCAKAKAPSSSDVMKKCLTRLKP 1620
DB 1561 FWEAVFTGLTHIDAFISQTOAGDNFAYLTAQATVCAKAKAPSSSDVMKKCLTRLKP 1620
QY 1621 TLVGPFTLLYRLSGVTNEVTLTHPPTKTIATCMQADLEVMTSTWVLAGVLAANAAYCIA 1680
DB 1621 TLVGPFTLLYRLSGVTNEVTLTHPPTKTIATCMQADLEVMTSTWVLAGVLAANAAYCIA 1680
QY 1681 TGCVCIIIGRLHINQRAVAVDPKEVLYEAFDEMEBCASPAALIIEGQRIAEMLKSKIQGL 1740
DB 1681 TGCVCIIIGRLHINQRAVAVDPKEVLYEAFDEMEBCASPAALIIEGQRIAEMLKSKIQGL 1740
QY 1741 QOASQAOADIQPTVOASWPKYEQFMAKMMNFISGIQYLAGLSTLPGNPAYASMAAFSA 1800


```

Db 1741 QOASKOADIOPAVOASWPKYEQFWAKKMMNFTSGIQYLAIGLSTLPGNPAVAMMAFSA 1800
Qy 1801 LTSPSTSTTLTLNLITGWLASQIAPBAGATGVVSGVGAAVSGIGKYLVDLACYG 1860
Db 1801 LTSPSTSTTLTLNLITGWLASQIAPBAGATGVVSGVGAAVSGIGKYLVDLACYG 1860
Qy 1861 AGISGALVAFKIMSGEKSPMEDVNLPLGILSPGALVVGIVCAALIRHVGEGEAVOMM 1920
Db 1861 AGISGALVAFKIMSGEKSPMEDVNLPLGILSPGALVVGIVCAALIRHVGEGEAVOMM 1920
Qy 1921 NRLIAFASRGNHVAPTHYVTESDASQRYTOLIGSLITLTLRLHNMITEDECPICGSM 1980
Db 1921 NRLIAFASRGNHVAPTHYVTESDASQRYTOLIGSLITLTLRLHNMITEDECPICGSM 1980
Qy 1981 LRDVDMWCTLLTBPCKMLTSLPKPMGLPVVSOCKYKGVWAGTGMTRPCGANIS 2040
Db 1981 LRDVDMWCTLLTBPCKMLTSLPKPMGLPVVSOCKYKGVWAGTGMTRPCGANIS 2040
Qy 2041 GNVRLGSMRTIGPKTCNNIWOCTPPINCYTEGQCVPKAPNPKVAILWVAASEYAEVQH 2100
Db 2041 GNVRLGSMRTIGPKTCNNIWOCTPPINCYTEGQCVPKAPNPKVAILWVAASEYAEVQH 2100
Qy 2101 GSYHYITGLTJTDNLKVPCLPSPPEFFSWVDGVQIHRFAPTPKPFPRDEVSFCVGLNSFV 2160
Db 2101 GSYHYITGLTJTDNLKVPCLPSPPEFFSWVDGVQIHRFAPTPKPFPRDEVSFCVGLNSFV 2160
Qy 2161 GSQLPCEDEPDTVLSMLTDPSSHITATTAARLRAGSPSPSSASSASQLSAPSLRAC 2220
Db 2161 GSQLPCEDEPDTVLSMLTDPSSHITATTAARLRAGSPSPSSASSASQLSAPSLRAC 2220
Qy 2221 THGKADVDVDMYDANLFMGDVTIRISGSKVVVLDLDMVEERSPLESIPEEYMLPKKR 2280
Db 2221 THGKADVDVDMYDANLFMGDVTIRISGSKVVVLDLDMVEERSPLESIPEEYMLPKKR 2280
Qy 2281 FPPALPAWARPDPNPLVESWKRDPYQATYAGALPPPRKTPPBPBRRTVGLSEDSI 2340
Db 2281 FPPALPAWARPDPNPLVESWKRDPYQATYAGALPPPRKTPPBPBRRTVGLSEDSI 2340
Qy 2341 GDALQOLAIFSGFQPPBSGDSGLSTGAGAADSGSTPPDELALESTGSISSMPLEGE 2400
Db 2341 GDALQOLAIFSGFQPPBSGDSGLSTGAGAADSGSTPPDELALESTGSISSMPLEGE 2400
Qy 2401 DPDLPEQVEQPPQGGVAAAPSGDSGSMSTCESEDDSVCCSMYSMTGALITCSPBE 2460
Db 2401 DPDLPEQVEQPPQGGVAAAPSGDSGSMSTCESEDDSVCCSMYSMTGALITCSPBE 2460
Qy 2461 EKLPIPLNSLRLRYHNKYCTTTSASLRAKVTFDPMQVLDYYSVLKDIKLASKV 2520
Db 2461 EKLPIPLNSLRLRYHNKYCTTTSASLRAKVTFDPMQVLDYYSVLKDIKLASKV 2520
Qy 2521 TARTLTMEACQLTPPHSARSKYFGAVNSLSGRAVNHISVWKDLEBSETPIPTTI 2580
Db 2521 TARTLTMEACQLTPPHSARSKYFGAVNSLSGRAVNHISVWKDLEBSETPIPTTI 2580
Qy 2581 TAKLTLTEACQLTPPHSARSKYFGAKEVNSLSGRAVNHISVWKDLEBSETPIPTTI 2640
Db 2581 TAKLTLTEACQLTPPHSARSKYFGAKEVNSLSGRAVNHISVWKDLEBSETPIPTTI 2640
Qy 2641 ORVEFLKAWAEKQDPMGFSYDTRCFDSTVTERDIRTESIYRACSLPEBATAHSLTE 2700
Db 2641 ORVEFLKAWAEKQDPMGFSYDTRCFDSTVTERDIRTESIYRACSLPEBATAHSLTE 2700
Qy 2701 RLHYGGPMFNKSGQTCYRRCASGVLTSMGNITTCYAKALAAKAGIAPTMLVCGD 2760
Db 2701 RLHYGGPMFNKSGQTCYRRCASGVLTSMGNITTCYAKALAAKAGIAPTMLVCGD 2760
Qy 2761 DLVVSSEQGTEDERNLRAFTEAMTRYAPGDPREPYDELITSCSSNVVALGPG 2820
Db 2761 DLVVSSEQGTEDERNLRAFTEAMTRYAPGDPREPYDELITSCSSNVVALGPG 2820
Qy 2821 RRRYVLTDPPTPIARAAMETVRHSVPVSWLGNIIQYPTIWARVNLTHFSSILMAQDT 2880
Db 2821 RRRYVLTDPPTPIARAAMETVRHSVPVSWLGNIIQYPTIWARVNLTHFSSILMAQDT 2880

```

```

Qy 2881 LDONLNFEMYGAVSVSPULDPAITERLHGDAFSLHTYTPHELTRVASALRKLGAPLR 2940
Db 2881 LDONLNFEMYGAVSVSPULDPAITERLHGDAFSLHTYTPHELTRVASALRKLGAPLR 2940
Qy 2941 AMKSRARAVRSLISRGRAAVCGRYLPNMAVKTUKLTPPEARLDDSSWFTVAGCG 3000
Db 2941 AMKSRARAVRSLISRGRAAVCGRYLPNMAVKTUKLTPPEARLDDSSWFTVAGCG 3000
Qy 3001 DIHYSRARBRLLLFGLLLFVGUGLFLPAR 3033
Db 3001 DIHYSRARBRLLLLSLLSVGVGLFLPAR 3033

RESULT 5
Q912A4
ID Q912A4 PRELIMINARY; PRT; 3033 AA.
AC Q912A4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MDa-4;
RA Itakura J., Nagayama K., Enomoto N., Kurosaki M., Watanabe H.,
RA Saco C.;
RU Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF238483; AAF59942.1; -.
DR HSPF; Q80YSL; ICW.
DR GO; GO:0016021; C: integral to membrane, IEA.
DR GO; GO:0019028; C: viral capsid, IEA.
DR GO; GO:0019031; C: viral envelope, IEA.
DR GO; GO:0005524; F: ATP binding, IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity, IEA.
DR GO; GO:0003723; F: RNA binding, IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity, IEA.
DR GO; GO:0008236; F: serine-type peptidase activity, IEA.
DR GO; GO:0005198; F: structural molecule activity, IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis, IEA.
DR GO; GO:0006350; P: transcription, IEA.
DR GO; GO:0019079; P: viral genome replication, IEA.
DR GO; GO:0019087; P: viral transformation, IEA.
DR InterPro; IPR000345; CYC_heme_BS.
DR InterPro; IPR001410; DEAD_BS.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RAP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept Ser Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_P5.
DR InterPro; IPR007094; RNA_pol_PSVit.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.

```

DR Pfam: PF00271; Helicase_C. 1.
DR Pfam: PF00998; Viral_Rdrp. 1.
DR SMART: SM00487; DEXDC. 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3033 AA, 329421 MW, 58A8ECD220F4D5CB CRC64;

Query Match 95.3%; Score 15356; DB 2; Length 3033;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 2864; Conservative 81; Mismatches 88; Indels 0; Gaps 0;

QY 1 MSNPKRQKTKNTRRPPDVKFGCGQIVGVYLLPRRGPLGTRARTRKTSKSPRG 60
DB 1 MSNPKRQKTKNTRRPPDVKFGCGQIVGVYLLPRRGPLGTRARTRKTSKSPRG 60
QY 61 RROPTRKDRSTKSGKPGYPMPLVGNBGLMAGMLSPRGSRPSMGNDPRHRNNG 120
DB 61 RROPTRKDRSTKSGKPGYPMPLVGNBGLMAGMLSPRGSRPSMGNDPRHRNNG 120
QY 121 KVIDLTGCFADLMGYIPVVGAPLGGVAPALAHGVLEBDGVNATGNLPGCSFSLFLA 180
DB 121 KVIDLTGCFADLMGYIPVVGAPLGGVAPALAHGVLEBDGVNATGNLPGCSFSLFLA 180
QY 181 LLSCTITPVSAAVKNI STGYMTNCTNDSTWOLQAALVHVPCCVPCCKGNASQCM 240
DB 181 LLSCTITPVSAAVKNI STGYMTNCTNDSTWOLQAALVHVPCCVPCCKGNASQCM 240
QY 241 PVSANVAORPGLTQGLRTHIDMVVMSATLCSALVYGDLCGGVMLAAMFTVSPQHMF 300
DB 241 PVSANVAORPGLTQGLRTHIDMVVMSATLCSALVYGDLCGGVMLAAMFTVSPQHMF 300
QY 301 VQDCNSIYEGITIGHRAMDMNNMNSPATMILAYAMRVEYIIDISGAHWGFMGLA 360
DB 301 VQDCNSIYEGITIGHRAMDMNNMNSPATMILAYAMRVEYIIDISGAHWGFMGLA 360
QY 361 YFSMOGAMAKVVIILLAGVDARHTVGGSAQTGRLTSIFDMGPRKIQLVNTNGW 420
DB 361 YFSMOGAMAKVVIILLAGVDANTY SVGSTAHLTQTFAGLFTLSQONIQLVNTNGW 420
QY 421 HIRKTLNCDNSLHTGFIALSFYTHSPNSGCEPERSACRSIEAPFVGMALQYEDNVN 480
DB 421 HIRKTLNCDNSLHTGFIALSFYTHSPNSGCEPERSACRSIEAPFVGMALQYEDNVN 480
QY 481 PEDMRPYCMHYPRQCGVNSAKTVCGPVYCTPSPVVVGTTDLGAPTYWGNETDVL 540
DB 481 PEDMRPYCMHYPRQCGVNSAKTVCGPVYCTPSPVVVGTTDLGAPTYWGNETDVL 540
QY 541 LNSTRPPLGSMFGCTWMNSSGYTKTCGAPPCRTRADFNASTDILCTDGFRAKHPDTYK 600
DB 541 LNSTRPPLGSMFGCTWMNSSGYTKTCGAPPCRTRADFNASTDILCTDGFRAKHPDTYK 600
QY 601 CGSGPMLTPRCLDYPRLMHVPCYNTYTFKIRMVVGVEHRLTAACNFTGDRCNLMD 660
DB 601 CGSGPMLTPRCLDYPRLMHVPCYNTYTFKIRMVVGVEHRLTAACNFTGDRCNLMD 660
QY 661 RDRSOLSPULHSTTEWAILPCSYDLPALSTGLHLHONIVDQVNYGLSPALTKYIVW 720
DB 661 RDRSOLSPULHSTTEWAILPCSYDLPALSTGLHLHONIVDQVNYGLSPALTKYIVW 720
QY 721 EWWIILFLLLADARVACMLILGQAEALAEKLVILHAASASNGELFYVIFVAVM 780
DB 721 EWWIILFLLLADARVACMLILGQAEALAEKLVILHAASASNGELFYVIFVAVM 780
QY 781 YIKGRAVPLATYSLTGLMSFSLLLALPOAYAYDASVHGQIGALLNMITLFTLPGYK 840
DB 781 YIKGRAVPLATYSLTGLMSFSLLLALPOAYAYDASVHGQIGALLNMITLFTLPGYK 840
QY 841 TLLSRELMLCYLLLTGEMAVQEMAPPMQYRGGRDGIWAVAFYGVGVFDITKMLAVL 900
DB 841 TLLSRELMLCYLLLTGEMAVQEMAPPMQYRGGRDGIWAVAFYGVGVFDITKMLAVL 900
QY 901 GRAYLLKGLATRPYFVRALALLRMCTMARHLAAGRYVOMALLALGRWTGYIYDHLTEM 960

DB 901 GPGYLRLGALTRVPYFVAHALLRMCTMARHLAAGRYVOMALLALGRWTGYIYDHLTEM 960
QY 961 SDMAASGRDLAVANVEPIIFSPMEKKVIVMGAEPTACGDIIHGHPVSARLREVLGPD 1020
DB 961 SDMAASGRDLAVANVEPIIFSPMEKKVIVMGAEPTACGDIIHGHPVSARLREVLGPD 1020
QY 1021 GYTSKWSLAPITAYAOQTGRLGTLIVVSMTGDKTBQAGEIOVLSTVTGSPFGTSISG 1080
DB 1021 GYTSKWSLAPITAYAOQTGRLGTLIVVSMTGDKTBQAGEIOVLSTVTGSPFGTSISG 1080
QY 1081 VLMTYTHAGNKITLAGSGPVTQWYSSABGDLVCHPSRPGTKSLPEPTCCGAVDLVLYRN 1140
DB 1081 VLMTYTHAGNKITLAGSGPVTQWYSSABGDLVCHPSRPGTKSLPEPTCCGAVDLVLYRN 1140
QY 1141 ADVIPARRGDKRGALSPRELSTLKSGSGPVLCPRHAGVFPAAVCSGVAKSIDFI 1200
DB 1141 ADVIPARRGDKRGALSPRELSTLKSGSGPVLCPRHAGVFPAAVCSGVAKSIDFI 1200
QY 1201 PVETLDIVTRSPFTSDNSTPPAVPQTVGYLHAPTGSGSKTVPAVAAQGYKVLVLP 1260
DB 1201 PVETLDIVTRSPFTSDNSTPPAVPQTVGYLHAPTGSGSKTVPAVAAQGYKVLVLP 1260
QY 1261 SYAATLGFAGYLSRAHGINPNI RTGVRTVTGDSITYSTYKFLADGGCAGAYDIIICD 1320
DB 1261 SYAATLGFAGYLSRAHGINPNI RTGVRTVTGDSITYSTYKFLADGGCAGAYDIIICD 1320
QY 1321 ECHAVDSTIIIGIGVLDQAEAGVRLTVLATATPPGCVTTPHPIEEVALGQEGEIPFY 1380
DB 1321 ECHAVDSTIIIGIGVLDQAEAGVRLTVLATATPPGCVTTPHPIEEVALGQEGEIPFY 1380
QY 1381 GRAIPLSYIKGRHLIFCHSKKKCDELAALRGMLNSVAYYRGDLSVIPQGDVVVA 1440
DB 1381 GRAIPLSYIKGRHLIFCHSKKKCDELAALRGMLNSVAYYRGDLSVIPQGDVVVA 1440
QY 1441 TDALMTGTGDPDSYIDCNVAVTQVVDLSLDPFTITTIQVLPDQAVSSQRRGTGRRL 1500
DB 1441 TDALMTGTGDPDSYIDCNVAVTQVVDLSLDPFTITTIQVLPDQAVSSQRRGTGRRL 1500
QY 1501 GIYRVSTGGERASGMFDSVLCCEYDAGAAYEILPSEETVRLAYFNTPGLPVCOHDLE 1560
DB 1501 GIYRVSTGGERASGMFDSVLCCEYDAGAAYEILPSEETVRLAYFNTPGLPVCOHDLE 1560
QY 1561 FWEAVFTGLTHIDAFHLSQTKQAGSFYLYAVQATVCARAKAPPSWDIMMKCLTRLP 1620
DB 1561 FWEAVFTGLTHIDAFHLSQTKQAGSFYLYAVQATVCARAKAPPSWDIMMKCLTRLP 1620
QY 1621 TLVGPTRLLYRLGVTNVTILTHPYTKYIATCMQADLEVMSTVWLAGVLAAYVCLA 1680
DB 1621 TLVGPTRLLYRLGVTNVTILTHPYTKYIATCMQADLEVMSTVWLAGVLAAYVCLA 1680
QY 1681 TGCVCIIIGRLHINORAVVAPDKEVLYEAFDMEBCASPAALIEGQRIAEMLKSKIQGL 1740
DB 1681 TGCVCIIIGRLHINORAVVAPDKEVLYEAFDMEBCASPAALIEGQRIAEMLKSKIQGL 1740
QY 1741 QOASKOADIOPVQASWPKVEQFQWAKMMNFI SGIQYIARLSTLPGNPAVASMMAFSA 1800
DB 1741 QOASKOADIOPVQASWPKVEQFQWAKMMNFI SGIQYIARLSTLPGNPAVASMMAFSA 1800
QY 1801 LTPSLSTSTTLLNLTGMLASQIAPPAAGAFVVSGLVGAANGSIGLKVLIQVITLGYG 1860
DB 1801 LTPSLSTSTTLLNLTGMLASQIAPPAAGAFVVSGLVGAANGSIGLKVLIQVITLGYG 1860
QY 1861 AGISGALVAFKIMSGEKSPMEDVNLLEGIISPGALVGVCAALIRRHVPGEGAVQM 1920
DB 1861 AGISGALVAFKIMSGEKSPMEDVNLLEGIISPGALVGVCAALIRRHVPGEGAVQM 1920
QY 1921 NRLIAFASRGHVAFTHYVTESSAQVTOQLGSLTITSLRLHMTTEDCPICGGSW 1980
DB 1921 NRLIAFASRGHVAFTHYVTESSAQVTOQLGSLTITSLRLHMTTEDCPICGGSW 1980
QY 1981 LRDVDMWCTLLTDPKMLTSLKLPFKMGLFPVSGQKYKVMAGTGMITTRPCGANIS 2040

```
Db 1981 LRDVMDWVCTLLTDFKMWLTSKLPFKMPLPEISQCKYKGVMAAGTGMITTRCPGKANIS 2040
Qy 2041 GNVRLGSRNRTGPKTKCMMIMOGTTPINCTEGOCVCPKPAENFKVAIMRYAASEVAVTQH 2100
Db 2041 GNVRLGSRNRTGPKTKCMMIMOGTTPINCTEGOCVCPKPAENFKVAIMRYAASEVAVTQH 2100
Qy 2101 GSYHYITGIJTDNLKVPQCPSPPEFSSWVDGQIHRFAPTPPKPFRFDEVSFCVGLNSFPV 2160
Db 2101 GQSYVTGTGTTDNMLKVPQCPSPPEFSSWVDGQIHRFAPTPPKPFRFDEVSFCVGLNSFPV 2160
Qy 2161 GSQLEPCDEPDTVMISMLTDBSHITAEIARLARGLARGLSPASSASQSLASPLRATCT 2220
Db 2161 GSQLEPCDEPDTVMISMLTDBSHITAEIARLARGLARGLSPASSASQSLASPLRATCT 2220
Qy 2221 THGKAYDVMVDANLFMGSDVTRISGSKVVVLDLDPVVERSLPEIPSEWMLPKKR 2280
Db 2221 THGKAYDVMVDANLFMGSDVTRISGSKVVVLDLDPVVERSLPEIPSEWMLPKKR 2280
Qy 2281 PPPALPAMARPDYNPLVESWKRPDYOPATVAGCALPPPKTPTPPRRRRRTVGLSEDSI 2340
Db 2281 PPPALPAMARPDYNPLVESWKRPDYOPATVAGCALPPPKTPTPPRRRRRTVGLSEDSI 2340
Qy 2341 GDALQOLAIKSTFGQPPSGDSGLSTGAGAAGSGQTPPEDELSETGSISSMPLEBEG 2400
Db 2341 GDALQOLAIKSTFGQPPSGDSGLSTGAGAAGSGQTPPEDELSETGSISSMPLEBEG 2400
Qy 2401 DPDLPEQVEPQPPQGGVAAAPSGDSGWSCTSEBDDSVCCSMGYSMTGALITPCSPPE 2460
Db 2401 DPDLPEQVEPQPPQGGVAAAPSGDSGWSCTSEBDDSVCCSMGYSMTGALITPCSPPE 2460
Qy 2461 EKLPIPLNSILRLYHNKVCYCTTTSASLRAKKTVPDRMQLDYSYDVLKDIKLAASKV 2520
Db 2461 EKLPIPLNSILRLYHNKVCYCTTTSASLRAKKTVPDRMQLDYSYDVLKDIKLAASKV 2520
Qy 2521 TARTLTMEBAQCLTPPHSARSKYFGAIVNSLSGRAVNHKSKVKDLLEDSERP1PTTI 2580
Db 2521 TARTLTMEBAQCLTPPHSARSKYFGAIVNSLSGRAVNHKSKVKDLLEDSERP1PTTI 2580
Qy 2581 SARLTLTEBAQCLTPPHSARSKYFGAIVNSLSGRAVNHKSKVKDLLEDSERP1PTTI 2580
Db 2581 SARLTLTEBAQCLTPPHSARSKYFGAIVNSLSGRAVNHKSKVKDLLEDSERP1PTTI 2580
Qy 2581 MAKEVFCVDPPTKGGKKAARLIVPDIGVRCERKALVDIQLKQAVMGASVGFQYSPBA 2640
Db 2581 MAKEVFCVDPPTKGGKKAARLIVPDIGVRCERKALVDIQLKQAVMGASVGFQYSPBA 2640
Qy 2641 QRVFELKAMAEKKDPMGFSDYTRCFDSVTVERDIRTEESIYRACSLBEAHTAHSLTE 2700
Db 2641 QRVFELKAMAEKKDPMGFSDYTRCFDSVTVERDIRTEESIYRACSLBEAHTAHSLTE 2700
Qy 2701 RLIVGGPMNSKGGQCYRRCASGVLTTSMGNTITCYVAKALAAKAGIAPVPTMLVCGD 2760
Db 2701 RLIVGGPMNSKGGQCYRRCASGVLTTSMGNTITCYVAKALAAKAGIAPVPTMLVCGD 2760
Qy 2761 DLVVISSEQTEBEDERNLRAPTEAMTRYASAPGDPREPREYDELTTCSSSNVSVALLPQG 2820
Db 2761 DLVVISSEQTEBEDERNLRAPTEAMTRYASAPGDPREPREYDELTTCSSSNVSVALLPQG 2820
Qy 2821 RRRVYLTDPPTPIARAAMETVRSHPVNSWLNIIQYAPTIARWVNLTHFFSIIMADT 2880
Db 2821 RRRVYLTDPPTPIARAAMETVRSHPVNSWLNIIQYAPTIARWVNLTHFFSIIMADT 2880
Qy 2881 LDONINFEYMGAVSVSPDLPAIIEHLHGLDAFSLHTYTPHELTRVASALRKLGAPPLR 2940
Db 2881 LDONINFEYMGAVSVSPDLPAIIEHLHGLDAFSLHTYTPHELTRVASALRKLGAPPLR 2940
Qy 2941 AMKSRARAVRASLSRGGRAAACGRLYLFNMAVKTLLKLTPLPEARLLDLSMFWTGAAGG 3000
Db 2941 AMKSRARAVRASLSRGGRAAACGRLYLFNMAVKTLLKLTPLPEARLLDLSMFWTGAAGG 3000
Qy 3001 DIYHSVSRAPRLILGLILLFVGGLFLPAR 3033
Db 3001 DIYHSVSRAPRLILGLILLFVGGLFLPAR 3033
```

RESULT 6
Q912A5

```
ID Q912A5 PRELIMINARY; PRT; 3033 AA.
AC Q912A5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepcivirus.
CX NCB1_Taxid=11103;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=MD2a-2;
RA Itakura J., Nagayama K., Enomoto N., Kurosaki M., Watanabe H.,
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF238482; AAF59941.1; -.
DR HSPD; O8JYS1; 1CW.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; Cytochrome BS.
DR InterPro; IPR001410; DEAD_heme_BS.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_NS5a.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR009003; Peptidase S29.
DR InterPro; IPR002518; Peptidase NS2.
DR InterPro; IPR007095; RNA_pol_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; Viral_RdRp; 1.
DR SMART; SM00487; DEDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polypeptide; Transmembrane.
SQ SEQUENCE 3033 AA; 329099 MW; 74E61003AEC7816A CRC64;
```

```
Query Match 95.3%; Score 1535; DB 2; Length 3033;
Best Local Similarity 94.2%; Pred. No. 0;
Matches 2857; Conservative 83; Mismatches 93; Indels 0; Gaps 0;

Qy 1 MSTNPKPQKTKRNTNRNRPQVKKFPGGQIVGVYLLPRGPRIGVARTKTSERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRNRPQVKKFPGGQIVGVYLLPRGPRIGVARTKTSERSQPRG 60
Qy 61 RRQPIKXDRSRGKSGWCKPGYVWPLVYNGEIGWAGWLLSPGSRPSPWGPNDPRHRSRVG 120
Db 61 RRQPIKXDRSRGKSGWCKPGYVWPLVYNGEIGWAGWLLSPGSRPSPWGPNDPRHRSRVG 120
```

Db 61 RR0PIKDRRSTGKSGKPGYPMPLYNEGGLMAGMLSPRGSRPSCGPTDRHRSRNVG 120
QY 121 KIIDLTGCFADLMGYIPVVGAPLGGVARALAHGVRLVEDGYNPATGNLPGCSFSIFLLA 180
Db 121 KIIDLTGCFADLMGYIPVVGAPLGGVARALAHGVRLVEDGYNPATGNLPGCSFSIFLLA 180
QY 181 LLSCTITPVSAABEVKNISTGVWVNDCTNDSITWQLOAVLHVPGCVPERKVGNSQCI 240
Db 181 LLSCTITPVSAABEVKNISTGVWVNDCTNDSITWQLOAVLHVPGCVPERKVGNSQCI 240
QY 241 PUSPNVAVORPGALTOGLRTHIDMYVMASLCSALYVGLCGGWLAAOMFIVSQHMF 300
Db 241 PUSPNVAVORPGALTOGLRTHIDMYVMASLCSALYVGLCGGWLAAOMFIVSQHMF 300
QY 301 VODCNCSITPGTITGHRMAMDMMMNSPATMIILAAKRVPEVIIDIIISGAMGVFGLA 360
Db 301 VODCNCSITPGTITGHRMAMDMMMNSPATMIILAAKRVPEVIIDIIISGAMGVFGLA 360
QY 361 YFSMOGAMAKVVVILLLAGVDARTHTVGGSAAGTTGLTSLFDMGPROKIQLVNTNSG 420
Db 361 YFSMOGAMAKVVVILLLAGVDARTHTHTTSVAGRTTSGPAGIFTSGPKNIQIINTNSG 420
QY 421 HNNRTRALNCNDSLHTGFILSLFTYHSFNSSGCPERSACRSIEARVVGALQYEDNTN 480
Db 421 HNNRTRALNCNDSLHTGFILSLFTYHSFNSSGCPERSACRSIEARVVGALQYEDNTN 480
QY 481 PEDMPPYCMHYPPOCGVVSATKVGCPVYCTPSPVVGTTDLRLGAPTYMGENDVFL 540
Db 481 PEDMPPYCMHYPPOCGVVSATKVGCPVYCTPSPVVGTTDLRLGAPTYMGENDVFL 540
QY 541 LNSTRPBLSGWFCTWMNNSGYTKTCGAPPCRTRADFNASTDLCPCTDFRKHPPDYLYK 600
Db 541 LNSTRPBLSGWFCTWMNNSGYTKTCGAPPCRTRADFNASTDLCPCTDFRKHPPDYLYK 600
QY 601 CGSGPMLTRCLIDPYRLMHYPCYNYTIFKIRMYGVGEHRLTRACNFTGBRCNLED 660
Db 601 CGSGPMLTRCLIDPYRLMHYPCYNYTIFKIRMYGVGEHRLTRACNFTGBRCNLED 660
QY 661 RDRSOLSPLLHSTTEWAILPCSYDLPALSTGLHLHONIVDQVPMYGLSPALTYKIVYM 720
Db 661 RDRSOLSPLLHSTTEWAILPCSYDLPALSTGLHLHONIVDQVPMYGLSPALTYKIVYM 720
QY 721 EBNVILLFLILADARVACIMLMLILGOAEALAEKVLIIHAASASNGFLYFVIEFVAAM 780
Db 721 EBNVILLFLILADARVACIMLMLILGOAEALAEKVLIIHAASASNGFLYFVIEFVAAM 780
QY 781 YIKGAVNPLATYSLTGLMFSLLILALPOOAYAYDASVHGQIALIWLWITFTLTGPKX 840
Db 781 YIKGAVNPLAAYSSTGLMFPCLLLVLPQOAYAYDASVHGQIALIILITLFTLTGPKX 840
QY 841 TLLSRFLMWLCYLLTLGEMVQEMAPPMQVGRGDGIWAVALIFYPVVFDITKMLAVL 900
Db 841 TLLSRCLMWLCYLLTLGEMVQEMAPPMQVGRGDGIWAVALIFYPVVFDITKMLAVL 900
QY 901 GPAYILLKGLATVPYFVRHALLRMCTMARHLAGARYVOMALLALGRWTGYIYDHLTPM 960
Db 901 GPAYILLKGLATVPYFVRHALLRMCTMARHLAGARYVOMALLALGRWTGYIYDHLTPM 960
QY 961 SPMAASGLDLAVNAVERIIFSPMEKKVIYWGAEFTAACGDIILGLPVSARLAGEVLLGPD 1020
Db 961 SPMAASGLDLAVNAVERIIFSPMEKKVIYWGAEFTAACGDIILGLPVSARLAGEVLLGPD 1020
QY 1021 GYTSKMSLILAPITAYAOQTRGLGTIVYSGRDTBOAGEIOVLSTYOSFLGTSISG 1080
Db 1021 GYTSKMSLILAPITAYAOQTRGLGTIVYSGRDTBOAGEIOVLSTYOSFLGTSISG 1080
QY 1081 VLMTYVHGAAGNTLAGSRPVTQMYSSABGLVGMSPSPGTXSLBECTGADVLYLVTRN 1140
Db 1081 VLMTYVHGAAGNTLAGSRPVTQMYSSABGLVGMSPSPGTXSLBECTGADVLYLVTRN 1140
QY 1141 ADVITARRRGDKGALLSPRPSTLKSGSGRPYLCPRGHNAVGFRAAVCSRGVAKSIDFI 1200
Db 1141 ADVITARRRGDKGALLSPRPSTLKSGSGRPYLCPRGHNAVGFRAAVCSRGVAKSIDFI 1200

QY 1201 PVEITLIDVTRSPTSNDNSTPPAVPQTOYGYLHAPTGSGKSTKYVVAAYAGYKVLVLNP 1260
Db 1201 PVEITLIDVTRSPTSNDNSTPPAVPQTOYGYLHAPTGSGKSTKYVVAAYAGYKVLVLNP 1260
QY 1261 SVAATLGFAYLSKRAHGINPNIARTGVRTVTTGABITYSTYVKFPLADGGCAGAYDIIICD 1320
Db 1261 SVAATLGFAYLSKRAHGINPNIARTGVRTVTTGABITYSTYVKFPLADGGCAGAYDIIICD 1320
QY 1321 ECHAVDSTTIIGICTVLDQAATAVRLTVLATAPRPGSVTTPHNIEVVALGOGGEIPFY 1380
Db 1321 ECHAVDSTTIIGICTVLDQAATAVRLTVLATAPRPGSVTTPHNIEVVALGOGGEIPFY 1380
QY 1381 GRAIPLSYIKGRHLIFCHSKKKCDLAAALRGGLNSVAAYYRGLDVSIPTOGDVVVVA 1440
Db 1381 GRAIPLSHIKGRHLIFCHSKKKCDLAAALRGGLNSVAAYYRGLDVSIPTOGDVVVVA 1440
QY 1441 TDAIMTGYTGDPSVIDCNVAVTQVDFSLDPTFTTTOIVPODAVSRQRGRTRGRRL 1500
Db 1441 TDAIMTGYTGDPSVIDCNVAVTQVDFSLDPTFTTTOIVPODAVSRQRGRTRGRRL 1500
QY 1501 GIVRYVSTGERASGMPDSVULCECYDAGAAMYELTPSETTVRLRAYFNTPLPVCOOHLF 1560
Db 1501 GIVRYVSTGERASGMPDSVULCECYDAGAAMYELTPSETTVRLRAYFNTPLPVCOOHLF 1560
QY 1561 FWEAVFTGLTHIDAHFLSQTQSGENFAYLTAYOATVCARAKAPPPSMDVMKCLTRLKP 1620
Db 1561 FWEAVFTGLTHIDAHFLSQTQSGENFAYLTAYOATVCARAKAPPPSMDVMKCLTRLKP 1620
QY 1621 TLVGPPTLLYRLSGVNEVTLTHPVTKYIATQWQADLEBVMSTVWLAGVLAAYVCLA 1680
Db 1621 TLVGPPTLLYRLSGVNEVTLTHPVTKYIATQWQADLEBVMSTVWLAGVLAAYVCLA 1680
QY 1681 TGCVCITIGRLHINORAVAPRKEVLYEAFDMECSASPAALIEBGORIAEMLKXKIQGL 1740
Db 1681 TGCVCITIGRLHINORAVAPRKEVLYEAFDMECSASPAALIEBGORIAEMLKXKIQGL 1740
QY 1741 QOASKOADIQPTVQASWPKEQFWAKAMNFIQIYTLAQLSTLPGNPAVAMMAFSA 1800
Db 1741 QOASKOADIQPTVQASWPKEQFWAKAMNFIQIYTLAQLSTLPGNPAVAMMAFSA 1800
QY 1801 LTPSLSTTILLNILGGLWLASQIAPAGATGVVSGLVGAASIGLGVLYDLIAGYG 1860
Db 1801 LTPSLSTTILLNILGGLWLASQIAPAGATGVVSGLVGAASIGLGVLYDLIAGYG 1860
QY 1861 AGISGALVAFKIMGEXKPSMEDVNLPGIILSPALVGVCAALIRRHVPGGAGVOM 1920
Db 1861 AGISGALVAFKIMGEXKPSMEDVNLPGIILSPALVGVCAALIRRHVPGGAGVOM 1920
QY 1921 NRLIAFASRGHVAPTHVYTESDASQVTOQLGSLITISLRLHMTTEDCPIPCGGSW 1980
Db 1921 NRLIAFASRGHVAPTHVYTESDASQVTOQLGSLITISLRLHMTTEDCPIPCGGSW 1980
QY 1981 LRDYDWCYCTLLDPKMWLSKLPKMGFLPFISQOKYKVMAGTGIMTRCCGANIS 2040
Db 1981 LRDYDWCYCTLLDPKMWLSKLPKMGFLPFISQOKYKVMAGTGIMTRCCGANIS 2040
QY 2041 GNVRLGSRITIGPTCKMNIWGTPEINCYTEGOCVPKAPFKAALMRVAASEVAVTQH 2100
Db 2041 GNVRLGSRITIGPTCKMNIWGTPEINCYTEGOCVPKAPFKAALMRVAASEVAVTQH 2100
QY 2101 GSYHYITGLTJTDNLKVPQOLSPSEFFSVWVDVQIHRFAPTPKPPFRDEVSCFVGLNSFV 2160
Db 2101 GSYHYITGLTJTDNLKVPQOLSPSEFFSVWVDVQIHRFAPTPKPPFRDEVSCFVGLNSFV 2160
QY 2161 GSQLPCEBEPDVLVMSMLTDPSHITAEARLARLARGSPSEASASSASQLSAPSLRATCT 2220
Db 2161 GSQLPCEBEPDVLVMSMLTDPSHITAEARLARLARGSPSEASASSASQLSAPSLRATCT 2220
QY 2221 THGKAYDVMDVDALEPMGQDVTRIESGSKYVVLDSLDPMEBERDLSEPISEYMLPKKR 2280
Db 2221 THGKAYDVMDVDALEPMGQDVTRIESGSKYVVLDSLDPMEBERDLSEPISEYMLPKKR 2280

```

QY 2281 FPPALPAMAPDPYNPLVESWKRPPYQATVAGCALPPRPKPTPTPPRRRRRTVGLSEBDSI 2340
DB 2281 FPPALPAMAPDPYNPLVESWKRPPYQATVAGCALPPRPKPTPTPPRRRRRTVGLSEBDSI 2340
QY 2341 GDALQOLAIKSGFGPPSGDSGLSTGAGAADSGSQTPPELALSETGSISSMPPLLEGELG 2400
DB 2341 GDALQOLAIKSGFGPPSGDSGLSTGAGAADSGSPPELALSETGSISSMPPLLEGELG 2400
QY 2401 DDLDEPEQVEPQPPQGGVAAAGSGSDSGSTSEBDDSVCCSMGYSWGTALITCSPPE 2460
DB 2401 DDLDEPEQVEPQPPQGGVAAAGSGSDSGSTSEBDDSVCCSMGYSWGTALITCSPPE 2460
QY 2461 EKLPIPNPLSNSLRHNKVCCTTTKSASLRARAKYTFDRMQVDSYDSDYKDIKLAASRY 2520
DB 2461 EKLPIPNPLSNSLRHNKVCCTTTKSASLRARAKYTFDRMQVDSYDSDYKDIKLAASRY 2520
QY 2521 TARRLTMBACQLTPPHSARSRYGFAKEVRSLSGRAVNHKISWKDLLEDSETP1PTTI 2580
DB 2521 TARRLTMBACQLTPPHSARSRYGFAKEVRSLSGRAVNHKISWKDLLEDSETP1PTTI 2580
QY 2581 MAKNEVFCVDPPTKGGKKAARLLVYPDLCVRCCEKALYDTQKLPQAVMGASGYFOYSPA 2640
DB 2581 MAKNEVFCVDPPTKGGKKAARLLVYPDLCVRCCEKALYDTQKLPQAVMGASGYFOYSPA 2640
QY 2641 QVBEFLKAMAEKQDPMGFSYDTRCFDSTVTERDRTREESIYRACSLPEAHATAHSLTE 2700
DB 2641 QVBEFLKAMAEKQDPMGFSYDTRCFDSTVTERDRTREESIYRACSLPEAHATAHSLTE 2700
QY 2701 RLTVGGMFNSKQTCGRRCRASGVLTTSMGNTTTCYKALAAKAGIAPMLVCGD 2760
DB 2701 RLTVGGMFNSKQTCGRRCRASGVLTTSMGNTTTCYKALAAKAGIAPMLVCGD 2760
QY 2761 DLVVISSESGCTEDERNLRAFTEAMTRYSAPRGDPPRPYDELITTCSSSNVVALGPGG 2820
DB 2761 DLVVISSESGCTEDERNLRAFTEAMTRYSAPRGDPPRPYDELITTCSSSNVVALGPGG 2820
QY 2821 RRRYLTLDPTTPIRAAMETVRSHPVNSWMLNIIQVATYARAVLMTMHPFSILMAOFT 2880
DB 2821 RRRYLTLDPTTPIRAAMETVRSHPVNSWMLNIIQVATYARAVLMTMHPFSILMAOFT 2880
QY 2881 LDQNLNFMFYGAIVSVSPDLPAIERLHGLDAFSLHTYTPHELTRVASALRKLGAPELR 2940
DB 2881 LDQNLNFMFYGAIVSVSPDLPAIERLHGLDAFSLHTYTPHELTRVASALRKLGAPELR 2940
QY 2941 AMKSAARAVRASLISGRGAACGRYLFRWAVYTKLKLTPLEBALDLSFPTVAGAGG 3000
DB 2941 AMKSAARAVRASLISGRGAACGRYLFRWAVYTKLKLTPLEBALDLSFPTVAGAGG 3000
QY 3001 DIYHSVSRARPLLLFGLLLFVGVGLFLLPAR 3033
DB 3001 DIYHSVSRARPLLLFGLLLFVGVGLFLLPAR 3033
QY 3001 HIYHSVSRARPLLLFGLLLFVGVGLFLLPAR 3033
DB 3001 HIYHSVSRARPLLLFGLLLFVGVGLFLLPAR 3033

RESULT 7
Q991B7 PRELIMINARY: PRT: 3033 AA.
AC Q991B7:
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DR Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepcivirinae.
OX NCB1_TaxID=11103;
RX [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21316767; PubMed=11424123; DOI=10.1002/jmv.1055;
RA Kato T., Furusaka A., Miyamoto M., Date T., Yasui K., Hiramoto J.,
RA Nagayama K., Tanaka T., Makita C.
RT "Sequence analysis of hepatitis C virus isolated from a fulminant
RT hepatitis patient."
RL J. Med. Virol. 64:334-339 (2001).

```

```

DR EMBL: AB047640; BAB32873.1; -.
DR HSSP: 08JYSL; 1CMX.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0019028; C: viral capsid; IEA.
DR GO: GO:0019031; C: viral envelope; IEA.
DR GO: GO:0005524; F: ATP binding; IEA.
DR GO: GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO: GO:0003723; F: RNA binding; IEA.
DR GO: GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO: GO:0008236; F: serine-type peptidase activity; IEA.
DR GO: GO:0005198; F: structural molecule activity; IEA.
DR GO: GO:0006350; P: transcription; IEA.
DR GO: GO:0019079; P: viral genome replication; IEA.
DR GO: GO:0019087; P: viral transformation; IEA.
DR InterPro: IPR000345; CyC_heme_BS.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR011545; DEAD/DEAH N.
DR InterPro: IPR002522; HCV capsid.
DR InterPro: IPR002521; HCV core.
DR InterPro: IPR002519; HCV env.
DR InterPro: IPR002531; HCV NS1.
DR InterPro: IPR000745; HCV NS4a.
DR InterPro: IPR001490; HCV NS4b.
DR InterPro: IPR002868; HCV NS5a.
DR InterPro: IPR002166; HCV RdRp.
DR InterPro: IPR004109; Peptidase S29.
DR InterPro: IPR009003; Peptidase Ser_Cys.
DR InterPro: IPR002518; Peptidase U39_HCV NS2.
DR InterPro: IPR007095; RNA pol_DS_PS.
DR InterPro: IPR007094; RNA pol_PsVlt.
DR Pfam: PF01543; HCV capsid.
DR Pfam: PF01542; HCV core1.
DR Pfam: PF01539; HCV env.
DR Pfam: PF01560; HCV NS1.
DR Pfam: PF01538; HCV NS2.
DR Pfam: PF02907; HCV NS3.
DR Pfam: PF01006; HCV NS4a.
DR Pfam: PF01001; HCV NS4b.
DR Pfam: PF01506; HCV NS5a.
DR Pfam: PF00998; Viral_RdRp.
DR SMART: SM00487; DEXDC.
DR PROSITE: PS00190; CYTOCHROME C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polypeptide; Transmembrane.
SQ SEQUENCE 3033 AA; 329035 MW; 74BD3B786F7D77D8 CRC64;

Query Match
Beat Local Similarity 94.4%; Score 1534; DB 2; Length 3033;
Matches 2863; Conservative 73; Mismatches 97; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRPPDVKPPGGQIVGVYLLPRGRGLGVRATRKTSERSOPRG 60
DB 1 MSTNPKPQKTKRNTNRPPDVKPPGGQIVGVYLLPRGRGLGVRATRKTSERSOPRG 60
QY 61 RROP1PKDRBSTGKSGKRGYPWPLVGNEGIGMAGMLSPGSRSPSGWPNDRHRSRYNG 120
DB 61 RROP1PKDRBSTGKSGKRGYPWPLVGNEGIGMAGMLSPGSRSPSGWPNDRHRSRYNG 120
QY 121 KVIDLTLCGFPADLMGYIVVGAFLGVARALAHGVRVLEDDGNFATGNLPGCSFSIFLLA 180
DB 121 KVIDLTLCGFPADLMGYIVVGAFLGVARALAHGVRVLEDDGNFATGNLPGCSFSIFLLA 180
QY 121 KVIDLTLCGFPADLMGYIVVGAFLGVARALAHGVRVLEDDGNFATGNLPGCSFSIFLLA 180
DB 121 KVIDLTLCGFPADLMGYIVVGAFLGVARALAHGVRVLEDDGNFATGNLPGCSFSIFLLA 180
QY 181 LISCITTPVSAAEVKNISITGMVNDCTNDSITWQLOAAVYLVHPCGVCCEKYGNAQOCI 240
DB 181 LISCITTPVSAAEVKNISITGMVNDCTNDSITWQLOAAVYLVHPCGVCCEKYGNAQOCI 240
QY 241 PVSBNVAQVRGALTOGRTHIIDVWVYATCSALYVGDLCGVWLAOMPTVSPQHMF 300
DB 241 PVSBNVAQVRGALTOGRTHIIDVWVYATCSALYVGDLCGVWLAOMPTVSPQHMF 300
QY 301 VQDCNCSITPITIGHRAMDMNMWNSPTATMILAYARVBEVLIIDISGAWGVMPGLA 360
DB 301 VQDCNCSITPITIGHRAMDMNMWNSPTATMILAYARVBEVLIIDISGAWGVMPGLA 360

```

Db 301 VQBCNCSIIYPGALTGHMMADMNMNSPTTMTIAVVRVBEVIDIISGAWGVMEGLA 360
Qy 361 YFSNQGAMAKVYVILLIAGVDARTHTVGSAAQTTGRLTSLFDMGPRKIQIUNTSW 420
Db 361 YFSNQGAMAKVYVILLIAGVDARTHTTGSAAQTTSSLASAFSPGARQIOLINTNSW 420
Qy 421 HINFTALNCDSLHTGTIASLPTHSHNSGCCPERMSACSIEAFVWGALOYEDNVN 480
Db 421 HINFTALNCDSLHTGTFTALFYIHKENSSGCEPERLSACRNIIEFRIGMALQYDNNVN 480
Qy 481 PEDRPVCMHPPROQCVSAKTYCGVYCTFSPPVVVGTJDLGAPTYWGENEDVFL 540
Db 481 PEDRPVCMHPPROQCVPAQIVCGFYCTFSPPVVVGTJDLGAPTYWGENEDVFL 540
Qy 541 LNSTRPPLSGMFGCTMNNSSGYTKTCGAPPCRTFADFNASTDILCPTDCFRKAPDTYTK 600
Db 541 LNSTRPSPSGMFGCTMNNSTGFTKTCGAPPCRTFADFNSTDLCPDTCFRKAPDTYTK 600
Qy 601 CGSGPMLTPKCLIDYPRILMHYCTVNYTIFKIRMYGVGEHRLTAACNTRDRCNLED 660
Db 601 CGSGPMLTPKCLVDYPRILMHYCTVNYSTFKIRMYGVGEHRLTAACNTRDRCNLED 660
Qy 661 RDRSOLPLHSTTEMALPCSYSDLPALSTGLLHQNIVDQPMYGSALTKYIYVR 720
Db 661 RDRSOLPLHSTTEMALPCSYSDLPALSTGLLHQNIVDQPMYGSALTKYIYVR 720
Qy 721 EWWILFLFLADARCACTMMLILGQAEALBKLVLHAASAASCNGLFYFVIFEVAA 780
Db 721 EWWILFLFLADARCACTMMLILGQAEALBKLVLHAASAASCNGLFYFVIFEVAA 780
Qy 781 YIKGRVPLATYISLTGMSFSLILALPOQAYAYDASVHQIGALLMILTTLTPGX 840
Db 781 YIKGRVPLATYISLTGMSFSLILALPOQAYAYDASVHQIGALLMILTTLTPGX 840
Qy 841 TLLSRFLMMLCYLLTLGEAMYQEWAPPMOYRGSGDGIYMAVAIFYGCVVDIKMLAV 900
Db 841 TLLSRFLMMLCYLLTLGEAMYQEWAPPMOYRGSGDGIYMAVAIFYGCVVDIKMLAV 900
Qy 901 GPAYLLKALTRVYFVRHAHLRMCTMAHLAGRYVQWALLALGRMTGYIYDLHTM 960
Db 901 GPAYLLKALTRVYFVRHAHLRMCTMAHLAGRYVQWALLALGRMTGYIYDLHTM 960
Qy 961 SDMAASGLRDLAAVEPIIFSPMEKKYIYWGAEFLAACGDLHLGPPSARIGREVLLGPD 1020
Db 961 SDMAASGLRDLAAVEPIIFSPMEKKYIYWGAEFLAACGDLHLGPPSARIGREVLLGPD 1020
Qy 1021 GYTSKGSMLAPITAYAQOTRGLIGTIYVSMGTGDKTEQAGEIYQVSTVYQSLGTSISG 1080
Db 1021 GYTSKGSMLAPITAYAQOTRGLIGTIYVSMGTGDKTEQAGEIYQVSTVYQSLGTSISG 1080
Qy 1081 VLMVTYHAGNKTLAGSRGPVQWYSSAEGDLVGMPSPGTKSLBPTCGADVLYVTN 1140
Db 1081 VLMVTYHAGNKTLAGSRGPVQWYSSAEGDLVGMPSPGTKSLBPTCGADVLYVTN 1140
Qy 1141 ADVIPARRBDKXGALLSPRLSTLKSSSGCVLCPRGHNAVGFRAAVCSRGVAKSIDP 1200
Db 1141 ADVIPARRBDKXGALLSPRLSTLKSSSGCVLCPRGHNAVGFRAAVCSRGVAKSIDP 1200
Qy 1201 PVETLIDVTRSPFSDNSTPAPVQTYQVGYLHAPTSKSTKPYAVAAQGVLYVLP 1260
Db 1201 PVETLIDVTRSPFSDNSTPAPVQTYQVGYLHAPTSKSTKPYAVAAQGVLYVLP 1260
Qy 1261 SVAATLFGAYVLSKAHGINPINIRTVTTGAPITVSTYKFLADOGCAGAYDIIICD 1320
Db 1261 SVAATLFGAYVLSKAHGINPINIRTVTTGAPITVSTYKFLADOGCAGAYDIIICD 1320
Qy 1321 ECHAVDSTTLIGTIVLDOAETAGVRLTVLATAIPBSVTTTPHPIEBVALGEGEIPRY 1380
Db 1321 ECHAVDSTTLIGTIVLDOAETAGVRLTVLATAIPBSVTTTPHPIEBVALGEGEIPRY 1380
Qy 1381 GRAIPLSYIKGGRHLIFCHSKKKCDLAAALRGKUNSVAYYRGLDVSVIPTQGDVVVVA 1440
Db 1381 GRAIPLSYIKGGRHLIFCHSKKKCDLAAALRGKUNSVAYYRGLDVSVIPTQGDVVVVA 1440

Qy 1441 TDLMTGYTDPFDSVIDCNVAVTQVVDPSLDPFTFTTQIVFQDAYSRSORGRTRGR 1500
Db 1441 TDLMTGYTDPFDSVIDCNVAVTQVVDPSLDPFTFTTQIVFQDAYSRSORGRTRGR 1500
Qy 1501 GIYRYVSGERASGMFDSVULCECYDAGAMELTTPSBTTVRLRAYFNTPGLPVQCDBE 1560
Db 1501 GIYRYVSGERASGMFDSVULCECYDAGAMELTTPSBTTVRLRAYFNTPGLPVQCDBE 1560
Qy 1561 FWEAVFTGLTHIDAFLLSOTQSGENFAYLTAQVATCARAKAPPSPNDVMMKILTRKP 1620
Db 1561 FWEAVFTGLTHIDAFLLSOTQSGENFAYLTAQVATCARAKAPPSPNDVMMKILTRKP 1620
Qy 1621 TLVGPPTLLYRLGSVTNEVTLTHPYTKYIATCMQADLEVMSTWYLAGVLAAYACIA 1680
Db 1621 TLVGPPTLLYRLGSVTNEVTLTHPYTKYIATCMQADLEVMSTWYLAGVLAAYACIA 1680
Qy 1681 TGCVCITIGRLHINQRAVAPDKEVYEAFFDEMEBCASPAALIEBQRIAEMLKSKIQL 1740
Db 1681 TGCVCITIGRLHINQRAVAPDKEVYEAFFDEMEBCASPAALIEBQRIAEMLKSKIQL 1740
Qy 1741 QOASQOADIOPVQASMPKVEQFWAKKMMNFIISGIOYLAGISTLPGNPAVAMMAFSA 1800
Db 1741 QOASQOADIOPVQASMPKVEQFWAKKMMNFIISGIOYLAGISTLPGNPAVAMMAFSA 1800
Qy 1801 LTPSLSTSTTLILNLLGMLASQIAPPAAGTGFVVSGLVGAAGSIGLKVLDIILAGY 1860
Db 1801 LTPSLSTSTTLILNLLGMLASQIAPPAAGTGFVVSGLVGAAGSIGLKVLDIILAGY 1860
Qy 1861 AGISGALVAFKIMSEKSPMEDVNNLPGIISPGALVGVCAALIRRHVPGGSAVOM 1920
Db 1861 AGISGALVAFKIMSEKSPMEDVNNLPGIISPGALVGVCAALIRRHVPGGSAVOM 1920
Qy 1921 NRIIPASRGNNVATHTVYTESDASQRTOLLSITSLRRLHNMTEEDCPIPCGGSW 1980
Db 1921 NRIIPASRGNNVATHTVYTESDASQRTOLLSITSLRRLHNMTEEDCPIPCGGSW 1980
Qy 1981 LRDVMDWCTIITDCKMLTJSLFPKMPGLPFVSCQKGVKVMAGTGIMTRCPGANIS 2040
Db 1981 LRDVMDWCTIITDCKMLTJSLFPKMPGLPFVSCQKGVKVMAGTGIMTRCPGANIS 2040
Qy 2041 GNVRLGSMRITGPKCMNINQSTPPIYCYTSGQCYPRAPNFKALIMVAASEVAEYQH 2100
Db 2041 GNVRLGSMRITGPKCMNINQSTPPIYCYTSGQCYPRAPNFKALIMVAASEVAEYQH 2100
Qy 2101 GSYHYITGLTIDNMLKVPQCLSPBEPFVWDGQIHRPAPTPKPPRDEVSFCVGLNSFV 2160
Db 2101 GSYHYITGLTIDNMLKVPQCLSPBEPFVWDGQIHRPAPTPKPPRDEVSFCVGLNSFV 2160
Qy 2161 GSQLPCEBPDVLMMLTDPSHITAEARLARLSPSEASSASQLSAPSLRATCT 2220
Db 2161 GSQLPCEBPDVLMMLTDPSHITAEARLARLSPSEASSASQLSAPSLRATCT 2220
Qy 2221 THGRKAYDVMDANLPMGQDVTRIESSKRVVLDLDMVEERBDLESITSEYMLPKR 2280
Db 2221 THGRKAYDVMDANLPMGQDVTRIESSKRVVLDLDMVEERBDLESITSEYMLPKR 2280
Qy 2281 PPPALPAMARPDYNPLVESMKRPDYQATVAGCALPPRRKTPPPPRRRRTVGLSEDSI 2340
Db 2281 PPPALPAMARPDYNPLVESMKRPDYQATVAGCALPPRRKTPPPPRRRRTVGLSEDSI 2340
Qy 2341 GDALQOLAIKSPGQPPSGDSGLSTGAGADSGSQTPPDELALSETGSISSMPLEGLG 2400
Db 2341 GDALQOLAIKSPGQPPSGDSGLSTGAGADSGSQTPPDELALSETGSISSMPLEGLG 2400
Qy 2401 DPLDEPQVEQPPROGVAAPSGDSGSWSTCSEBDSVVCSSMSYMTGALLTPCSDEE 2460
Db 2401 DPLDEPQVEQPPROGVAAPSGDSGSWSTCSEBDSVVCSSMSYMTGALLTPCSDEE 2460
Qy 2461 EKLPINPISNSILRRHNVYCTTTSKASLRAKKVTFDMQVLDYSYDVLDOIKLAAKV 2520
Db 2461 EKLPINPISNSILRRHNVYCTTTSKASLRAKKVTFDMQVLDYSYDVLDOIKLAAKV 2520


```

QY 2521 TARTLTMEBAQOLTPPHSARSXYGFGAKEVRSLSGRAVNHKISVWMDLLEDEETPIPTTI 2580
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2521 SARLTLTEBAQOLTPPHSARSXYGFGAKVRSLSGAVNHHSVWMDLLEDEQTPPTTI 2580
QY 2581 MAKNEVFCVDPYKGGKKAARLIVYDPLGYRVEBKALYDTOKLPOAVMGASYGFOYSPA 2640
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2581 MAKNEVFCVDPYKGGKKAARLIVYDPLGYRVEBKALYDTOKLPOAVMGASYGFOYSPA 2640
QY 2641 QVVEFLTKMAAKKQPMGFSYDTRCDSTVTERDRIETESIRACSLPEEATAHLSLHE 2700
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2641 QVVEFLTKMAAKKQPMGFSYDTRCDSTVTERDRIETESIRACSLPEEATAHLSLHE 2700
QY 2701 RLYVGGPMFNSKQTCGYRRCASGVLTTSMGNTTTCYKALACACAGIAPTMVLCSD 2760
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2701 RLYVGGPMFNSKQTCGYRRCASGVLTTSMGNTTTCYKALACACAGIAPTMVLCSD 2760
QY 2761 DLVVISSEGTEDERNLRAFTAMTRYSAPPGDPPREYDLELITSCSNVSVALGPGQ 2820
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2761 DLVVISSEGTEDERNLRAFTAMTRYSAPPGDPPREYDLELITSCSNVSVALGPGQ 2820
QY 2821 RRRYLTDRPTPIRAAWEIVRHSVNSWLGNTIOYAPTIWARMVMTHEFSIIMAQDT 2880
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2821 RRRYLTDRPTPIRAAWEIVRHSVNSWLGNTIOYAPTIWARMVMTHEFSIIMAQDT 2880
QY 2881 LDONINFEWYGAIVSVSPDLPAIETRLHGLDAFSIHTYTPHELTRVASALRKLGPPLR 2940
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2881 LDONINFEWYGAIVSVSPDLPAIETRLHGLDAFSIHTYTPHELTRVASALRKLGPPLR 2940
QY 2941 AWKSRARAVRASLISRGGAACGRLTFWMAVTKTKLTPLEBARLLDLSWFTVAGAGG 3000
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2941 AWKSRARAVRASLISRGGAACGRLTFWMAVTKTKLTPLEBARLLDLSWFTVAGAGG 3000
QY 3001 DIYHSVSRARPRLILFGLLLTFVGVCLFLTPAR 3033
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3001 DIYHSVSRARPRLILFGLLLTFVGVCLFLTPAR 3033

```

RESULT 8

```

0991B3
ID 0991B3 PRELIMINARY; PRT: 3033 AA.
AC 0991B3;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxId=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21316767; PubMed=11424123; DOI=10.1002/jmv.1055;
RA Kato T., Furusaka A., Miyamoto M., Date T., Yasui K., Hiramoto J.,
RA Nagayama K., Tanaka T., Wakita T.;
RT "Sequence analysis of hepatitis C virus isolated from a fulminant
RT hepatitis patient.";
RL J. Med. Virol. 64:334-339 (2001).
DR EMBL; AB047644; BAB32877.1; -.
DR HSP; O8Y51; 1CWK.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0015079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; Cyt_c_heme_BS.

```

```

DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR007445; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS5a.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRp.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_Psitv.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; Viral_RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polypeptide; Transmembrane.
SQ SEQUENCE 3033 AA; 328908 MW; B306821483E70CB3 CRC64;

Query Match          95.2%  Score 15350;  DB 2;  Length 3033;
Best Local Similarity 94.3%  Pred No 0;
Matches 2860;  Conservative 81;  Mismatches 92;  Indels 0;  Gaps 0;

QY 1 MSTNPKPQKTKRNTNRBPQDVKFPFGGQIVGVYLLPRGRGLGVRAITRKTSESQPG 60
Db 1 MSTNPKPQKTKRNTNRBPQDVKFPFGGQIVGVYLLPRGRGLGVRAITRKTSESQPG 60
QY 61 RRQPIPKRRSTGSKGKRPYWPPLYXGNEGGLGMAWLLSPGSRSRSGPNDRHRSRVNG 120
Db 61 RRQPIPKRRSTGSKGKRPYWPPLYXGNEGGLGMAWLLSPGSRSRSGPNDRHRSRVNG 120
QY 121 KYIDLTGCPADLMGYIVVGAPLGGVABALAHGVRVLEDDGVNFPATGNLPGCSFSIFLLA 180
Db 121 KYIDLTGCPADLMGYIVVGAPLGGVABALAHGVRVLEDDGVNFPATGNLPGCSFSIFLLA 180
QY 181 LLSCTITPVSAAEVKNISGTWVNTDCTNDSITWQLQAAVLHVPQVPCBEKVGNSQCI 240
Db 181 LLSCTITPVSAAEVKNISGTWVNTDCTNDSITWQLQAAVLHVPQVPCBEKVGNSQCI 240
QY 241 PVSPPVAVQRCALTOGRLTHIDMYVMSATLCSALYVGDLCGVMWLAQMFIVSQHHWF 300
Db 241 PVSPPVAVQRCALTOGRLTHIDMYVMSATLCSALYVGDLCGVMWLAQMFIVSQHHWF 300
QY 301 VDDCNCSIYPGTTGTHRAMDMMMWSPATMTIILAYARVPEVITDIISGAMGVWFGLA 360
Db 301 VDDCNCSIYPGTTGTHRAMDMMMWSPATMTIILAYARVPEVITDIISGAMGVWFGLA 360
QY 361 YFSMGAAKAVVILLLAGVDARTHTVGGSAQAOTGRLTSLFDMPROKIQLVNTNGSW 420
Db 361 YFSMGAAKAVVILLLAGVDARTHTHTGAVAAASARGLASLFTSGPQNNQILNTNGSW 420
QY 421 HINRTALNCNDSLHTGFIASLFTYHSFNSSGCPREMSACRSIEARVVGALQYEDNTVN 480
Db 421 HINRTALNCNDSLHTGFIASLFTYHSFNSSGCPRELSVCRNIEARIGMGAQYEDNTVN 480
QY 481 PEDMSPYCMHPPEPCGVSASAKVGPYCCFTPSVWVGTTDRGLAPTYTWGENETDVEL 540
Db 481 PEDMSPYCMHPPEPCGVSASAKVGPYCCFTPSVWVGTTDRGLAPTYTWGENETDVEL 540
QY 541 INSTRPLGWSFGCTWNNSSGYTKTGAPPCRTADFNASTDLCLPTDCFRKHPTTYLK 600
Db 541 INSTRPLGWSFGCTWNNSSGYTKTGAPPCRTADFNASTDLCLPTDCFRKHPTTYLK 600

```

Db 541 LNSTRPBLSWFGCTWMNSTGFTKTCGAPCRIRADFNSTDDLCEPTDCFRKHDPATYIK 600

Qy 601 CGSGPMLTPRCLIDYPRYLWHYPCVNYTIFKIRMYGVGEHRLTAACNFTGDRCNLEB 660

Db 601 CGSGPMLTPRCLIDYPRYLWHYPCVNYTIFKIRMYGVGEHRLTAACNFTGDRCNLEB 660

Qy 661 RDRSQSLPLHSTTEWAILPCSYSDLPALSTGLLHONIVDVQPMYGLSPALTKYIVM 720

Db 661 RDRSQSLPLHSTTEWAILPCSYSDLPALSTGLLHONIVDVQPMYGLSPALTKYIVM 720

Qy 721 EWWILLFLILADARVACIMMLILGQAEALFKVILHAASASONGFLYPIFVAVAM 780

Db 721 EWWILLFLILADARVACIMMLILGQAEALFKVILHAASASONGFLYPIFVAVAM 780

Qy 781 YIKGRVPLATYSLTGLMSFSLILALPOQAVAYDASVHGQIGALLVMTLFTLTPGX 840

Db 781 YIKGRVPLATYSLTGLMSFSLILALPOQAVAYDASVHGQIGALLVMTLFTLTPGX 840

Qy 841 TLLSRFLMMLCYLLTIGEAHVQEMAPPMQVRGRDGIIMAAVAFYRGVVPDITKMLAVL 900

Db 841 TLLSRFLMMLCYLLTIGEAHVQEMAPPMQVRGRDGIIMAAVAFYRGVVPDITKMLAVL 900

Qy 901 GPAYLLKGALTVPYFVRAHALLRMCMTMAHLAGRHYVOMALLALGRWGTYYIDHLTPM 960

Db 901 GPAYLLRDALTVPYFVRAHALLRMCMTMAHLAGRHYVOMALLALGRWGTYYIDHLTPM 960

Qy 961 SDMAASGLRDIAVAVERIIFSPMEKVIYVGAETAAQDIIHGLPVSAIRGREVLLGPAD 1020

Db 961 SDMAASGLRDIAVAVERIIFSPMEKVIYVGAETAAQDIIHGLPVSAIRGREVLLGPAD 1020

Qy 1021 GYTSKQMSLIAPITAAVAAQOTRGLLGTIVSMTGRDKTEQAGEIOVLSTVTSQSLGTSISG 1080

Db 1021 GYTSKQMSLIAPITAAVAAQOTRGLLGTIVSMTGRDKTEQAGEIOVLSTVTSQSLGTSISG 1080

Qy 1081 VLMTVHVGAGNKTLAGSRGPVTOMSSABGDLGWSRPGCTKSLEBCTCGAVDLYVTMN 1140

Db 1081 VLMTVHVGAGNKTLAGSRGPVTOMSSABGDLGWSRPGCTKSLEBCTCGAVDLYVTMN 1140

Qy 1141 ADVIPARRRDKRGALSPRLSTLKGSSGGPVLCPRGHAVGVFRAAVCSRGVAKSIDT 1200

Db 1141 ADVIPARRRDKRGALSPRLSTLKGSSGGPVLCPRGHAVGVFRAAVCSRGVAKSIDT 1200

Qy 1201 PVETLIDIVTRSPTFSDNSTPRAVPOQYQVGYLHAPGSGKSTVPVAAVAAQYKVLVLP 1260

Db 1201 PVETLIDIVTRSPTFSDNSTPRAVPOQYQVGYLHAPGSGKSTVPVAAVAAQYKVLVLP 1260

Qy 1261 SVAAATIGFGAYLSKAGHININIRTVGRTVTTGAPITYSTYTGKRLAGCGAGAYDIIICD 1320

Db 1261 SVAAATIGFGAYLSKAGHININIRTVGRTVTTGAPITYSTYTGKRLAGCGAGAYDIIICD 1320

Qy 1321 ECHAANDSTTLIGITVLDQAEFAGVRLTVLATATPPGAVTTPHNTIEVALGQEGEIPRY 1380

Db 1321 ECHAANDSTTLIGITVLDQAEFAGVRLTVLATATPPGAVTTPHNTIEVALGQEGEIPRY 1380

Qy 1381 GRAIPLUSHIKGRHLIFCHSKKKCDELAALRGKGINAVAYYRGDLVSIIPFGDVAVVA 1440

Db 1381 GRAIPLUSHIKGRHLIFCHSKKKCDELAALRGKGINAVAYYRGDLVSIIPFGDVAVVA 1440

Qy 1441 TDALMTCYTGDPSVIDCNVAVTVQVDFSLDPTFTTTQIVPQDAVSRQRRTGRGL 1500

Db 1441 TDALMTCYTGDPSVIDCNVAVTVQVDFSLDPTFTTTQIVPQDAVSRQRRTGRGL 1500

Qy 1501 GIYRYVSTGBRAGMPSVVLCECYDAGAAWELTSEETVRLRAVFNTPGLPVCODHLE 1560

Db 1501 GIYRYVSTGBRAGMPSVVLCECYDAGAAWELTSEETVRLRAVFNTPGLPVCODHLE 1560

Qy 1561 FMEAVFTGLTHIDAHFLISQTKSGENFAVYLTAVOATVCAKAKAPPPSMDVMKCLTRLKP 1620

Db 1561 FMEAVFTGLTHIDAHFLISQTKSGENFAVYLTAVOATVCAKAKAPPPSMDVMKCLTRLKP 1620

Qy 1621 TLVGGPRLLYRLGSVYNEVTLTHPVTKYIATCMQADLEVTSTMVLGGVLAVAAYCILA 1680

Db 1621 TLVGGPRLLYRLGSVYNEVTLTHPVTKYIATCMQADLEVTSTMVLGGVLAVAAYCILA 1680

Qy 1681 TGCVCIIIGRLHINORAVVAPDKEVLYEAFDEMECASAALIEBGORIAEMLKSKIQGL 1740

Db 1681 TGCVCIIIGRLHINORAVVAPDKEVLYEAFDEMECASAALIEBGORIAEMLKSKIQGL 1740

Qy 1741 QOASKOADIOPTVOASWPKYEQFWAKMAMNFIGIOYLAQSLTFCGNPAVAMMAFSAA 1800

Db 1741 QOASKOADIOPTVOASWPKYEQFWAKMAMNFIGIOYLAQSLTFCGNPAVAMMAFSAA 1800

Qy 1801 LTPSLSTSTTLINILGGMWLSQIAPPAAGAGFVSGVGAAVSFIGKVLVHILAGYG 1860

Db 1801 LTPSLSTSTTLINILGGMWLSQIAPPAAGAGFVSGVGAAVSFIGKVLVHILAGYG 1860

Qy 1861 AGISGALVAFKIMSGEKSPMEDVNNLJPGILSPGALVGVICAAILRRHVPGEGAVOM 1920

Db 1861 AGISGALVAFKIMSGEKSPMEDVNNLJPGILSPGALVGVICAAILRRHVPGEGAVOM 1920

Qy 1921 NRIIAFASIGHNVAPTHYVTESDASQRYTQLLGSITISLRLRHNMTEDCPIPCAGSW 1980

Db 1921 NRIIAFASIGHNVAPTHYVTESDASQRYTQLLGSITISLRLRHNMTEDCPIPCAGSW 1980

Qy 1981 LRDVMDWCTILTPFKMWTSLKLPKMPGLPFVSGQKYGKVMAGTGMTRCCGANIS 2040

Db 1981 LRDVMDWCTILTPFKMWTSLKLPKMPGLPFVSGQKYGKVMAGTGMTRCCGANIS 2040

Qy 2041 GNVRLGSRITGPKTGMNIWQGTPEINCYTEGQCVPKPAFPKVAIMVAASEYAEVTOH 2100

Db 2041 GNVRLGSRITGPKTGMNIWQGTPEINCYTEGQCVPKPAFPKVAIMVAASEYAEVTOH 2100

Qy 2101 GSYHYITGLTTDNLKVPOLPSPPEFSSWDOVQIHRFAPTRKPPFRDVSCTVGLNSRV 2160

Db 2101 GSYHYITGLTTDNLKVPOLPSPPEFSSWDOVQIHRFAPTRKPPFRDVSCTVGLNSRV 2160

Qy 2161 GSQLPCEDEPPTDVMMSWLTJPSHITABTAARLARGSPSEBASASSAQSLAPSLRATCT 2220

Db 2161 GSQLPCEDEPPTDVMMSWLTJPSHITABTAARLARGSPSEBASASSAQSLAPSLRATCT 2220

Qy 2221 THGKAYDVDMYDANLFMGGDVYTRIESGSKVVVLDSLPMVERSDLEBSIPSEYMLPKR 2280

Db 2221 THGKAYDVDMYDANLFMGGDVYTRIESGSKVVVLDSLPMVERSDLEBSIPSEYMLPKR 2280

Qy 2281 FPPALPAMARPDYNPPLVESMKRPDYOPATYAGCALPPPKRTPTPPPPRRRTVGLSEDSI 2340

Db 2281 FPPALPAMARPDYNPPLVESMKRPDYOPATYAGCALPPPKRTPTPPPPRRRTVGLSEDSI 2340

Qy 2341 GDALQOLAIKFRGPPSGDSGLSTGADAADSGGKTPPDEBSALSTGSISSMPPLGEPG 2400

Db 2341 GDALQOLAIKFRGPPSGDSGLSTGADAADSGGKTPPDEBSALSTGSISSMPPLGEPG 2400

Qy 2401 DPDLPEQVEPOPPOGQVAAFGSDSGSWSTCSEEDBSVVCSSMSYWTGALITTPCSPEE 2460

Db 2401 DPDLPEQVEPOPPOGQVAAFGSDSGSWSTCSEEDBSVVCSSMSYWTGALITTPCSPEE 2460

Qy 2461 EKLPIPLNSLSLRHYNKVYCTTTKASLRAKCTTFFDMQVLDSYDSVLKOIKLAASKV 2520

Db 2461 EKLPIPLNSLSLRHYNKVYCTTTKASLRAKCTTFFDMQVLDSYDSVLKOIKLAASKV 2520

Qy 2521 TABRLTMBEACOLPPHARSKYGFAKEVNSLSGRAVNHKISVWKDLLEDSQPIPTTI 2580

Db 2521 TABRLTMBEACOLPPHARSKYGFAKEVNSLSGRAVNHKISVWKDLLEDSQPIPTTI 2580

Qy 2581 MAKNEVFCVDPPTKGKKAARLIYPPDLGVRCEKKAALYDITOKLPOAVMGASYGQYSPA 2640

Db 2581 MAKNEVFCVDPPTKGKKAARLIYPPDLGVRCEKKAALYDITOKLPOAVMGASYGQYSPA 2640

Qy 2641 QREVELLQAAABEKDPMGFSTIDTRCFDSTVTBERDITREBSIYRACSLPEEAHTAHSLTE 2700

Db 2641 QREVELLQAAABEKDPMGFSTIDTRCFDSTVTBERDITREBSIYRACSLPEEAHTAHSLTE 2700

Qy 2701 RLYVGGPMNSKGGOTCGYRCRASGVLTTSNGNTTTCVVKLAACKAGIILAPTMLVCGD 2760

Db 2701 RLYVGGPMNSKGGOTCGYRCRASGVLTTSNGNTTTCVVKLAACKAGIILAPTMLVCGD 2760

```

QY 2761 DLVVISSESGTEDEENLRAPTEAMTRYAPGDPPEPRPEYDELITSCSSNVSVALGPOG 2820
DB 2761 DLVVISSESGTEDEENLRAPTEAMTRYAPGDPPEPRPEYDELITSCSSNVSVALGPOG 2820
QY 2821 RRRYITRDPPTPIARAAMEVTRHSPVNSWLGNIIOYAPTIAWVLMTHPFSIIMAQDT 2880
DB 2821 RRRYITRDPPTPIARAAMEVTRHSPVNSWLGNIIOYAPTIAWVLMTHPFSIIMAQDT 2880
QY 2881 LDONINFEWYGVSVSPDLPIILERLHGLDAFSLHTTTPHELTPRVASALRKLGPPLR 2940
DB 2881 LDONINFEWYGVSVSPDLPIILERLHGLDAFSLHTTTPHELTPRVASALRKLGPPLR 2940
QY 2941 AKKSRAVARASLISGRRAAVCGRYLFNMWAVTKLKLTPLEARLLDSSMFTVGAGGG 3000
DB 2941 AKKSRAVARASLISGRRAAVCGRYLFNMWAVTKLKLTPLEARLLDSSMFTVGAGGG 3000
QY 3001 DIYHSVSRARPRLLFGLLLLFVGVGLFLLPAR 3033
DB 3001 DIYHSVSRARPRLLFGLLLLFVGVGLFLLPAR 3033

RESULT 9
QY12A6 PRELIMINARY; PRT; 3033 AA.
AC QY12A6;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepcivirus.
OX NCBI_Taxid=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MD2-1;
RA Itakura J., Nagayama K., Enomoto N., Kurosaki M., Watanabe H.,
RA Sato C.;
RU Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL, AF234481; AAF59940.1; -.
DR HSP; O8Y51; 1CWx.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; F:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro: IPR001345; Cyclic_heme_BS.
DR InterPro: IPR001410; DEAD_heme_BS.
DR InterPro: IPR011545; DEAD/DEAH_N.
DR InterPro: IPR002522; HCV_capsid.
DR InterPro: IPR002521; HCV_core.
DR InterPro: IPR002519; HCV_env.
DR InterPro: IPR002531; HCV_NS1.
DR InterPro: IPR000745; HCV_NS4a.
DR InterPro: IPR001490; HCV_NS4b.
DR InterPro: IPR002868; HCV_NS5a.
DR InterPro: IPR002166; HCV_NS5a.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR009003; Pept_Ser_Cys.
DR InterPro: IPR002518; Pept_U39_HCV_NS2.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.

```

```

DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; Viral_RdRp; 1.
DR SMART; SM00487; DEXDc; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3033 AA; 329445 MW; 9D0F544E06CA2B28 CRC64;

Query Match          95.2%; Score 15347; DB 2; Length 3033;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 2860; Conservative 82; Mismatches 91; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNNRRRPODYKPPGGQIYGVVLLPRRGPRGLGVARTRKTSERSQPRG 60
DB 1 MSTNPKPQKTKRNNRRRPODYKPPGGQIYGVVLLPRRGPRGLGVARTRKTSERSQPRG 60
QY 61 RROP1PKDORRSTGKSWGKPYWPLYGNEGGLGMAQMLSPRGSRPSWGPNDPRHRSRVNG 120
DB 61 RROP1PKDORRSTGKSWGKPYWPLYGNEGGLGMAQMLSPRGSRPSWGPNDPRHRSRVNG 120
QY 121 KYIDPLTGGFADLMGYIPYVGAFLGCVARALAHGVYLEDGVNFKTNLPGCSBSIFLLA 180
DB 121 KYIDPLTGGFADLMGYIPYVGAFLGCVARALAHGVYLEDGVNFKTNLPGCSBSIFLLA 180
QY 181 ILSCITTPVSAEAVNISTGVVNTDCTNDSITTWQLOAAVLHPGCVPECEKYNASQCVI 240
DB 181 ILSCITTPVSAEAVNISTGVVNTDCTNDSITTWQLOAAVLHPGCVPECEKYNASQCVI 240
QY 241 PVSFNVAVQRPQALTOGLRTHIDMVVMSATLCSALYVGLCGVMAAQMFTVSPQHMF 300
DB 241 PVSFNVAVQRPQALTOGLRTHIDMVVMSATLCSALYVGLCGVMAAQMFTVSPQHMF 300
QY 301 VODCNCSTYPTGTTTHRAAMDMMNWSPTATMTILANRVRPVITDIISGAMGVMFGIA 360
DB 301 VODCNCSTYPTGTTTHRAAMDMMNWSPTATMTILANRVRPVITDIISGAMGVMFGIA 360
QY 361 YFSMGAAKAVVILLLAGVDARHTTVGSAQOTGRULTSLPDMGPQKIQIOLVNTGSM 420
DB 361 YFSMGAAKAVVILLLAGVDARHTTVGSAQOTGRULTSLPDMGPQKIQIOLVNTGSM 420
QY 421 HINRTALNCNDSIHTGFIASLFYTHSFNSSCPERMSACRSIEARVGMGALQYEDNVTN 480
DB 421 HINRTALNCNDSIHTGFIASLFYTHSFNSSCPERMSACRSIEARVGMGALQYEDNVTN 480
QY 481 PEDMKPYCMHYPPKCGIVPARSVGYPVCFPPSVVVGGTDXGLVPYTWGENTDVFLL 540
DB 481 PEDMKPYCMHYPPKCGIVPARSVGYPVCFPPSVVVGGTDXGLVPYTWGENTDVFLL 540
QY 541 LNSTRPPLGSMFGCTMMNMSGYTKTCGAPPCQTRBDFNASTDILCPTCCFRKGPPTTYX 600
DB 541 LNSTRPPLGSMFGCTMMNMSGYTKTCGAPPCQTRBDFNASTDILCPTCCFRKGPPTTYX 600
QY 601 CGSGPWLTPRCLIDYVRLMHYPCVTNYTFKIRMYGVGEHRLTAACNFTGRDCNLED 660
DB 601 CGSGPWLTPRCLIDYVRLMHYPCVTNYTFKIRMYGVGEHRLTAACNFTGRDCNLED 660
QY 661 RDRSQSLPLHSTTSMALLPGSYSLPALSTGLHLHONIVDVQMYGSLPALTYRYVRW 720
DB 661 RDRSQSLPLHSTTSMALLPGSYSLPALSTGLHLHONIVDVQMYGSLPALTYRYVRW 720
QY 721 EKVILLFLLLDARVACACMMMLILGQAEALAEKLVILHAASASACNGLFYVIFPVAAW 780
DB 721 EKVILLFLLLDARVACACMMMLILGQAEALAEKLVILHAASASACNGLFYVIFPVAAW 780
QY 781 YIKGRVPLATYSLTGLWSFSLILALPOQAAVAVDASVHGQIGALLVMITLFTLLTPGYK 840
DB 781 YIKGRVPLATYSLTGLWSFSLILALPOQAAVAVDASVHGQIGALLVMITLFTLLTPGYK 840

```

Db 781 YIKRAVPLAAVSLTGLMPCLLVLA.PQOAVAYDASVHGQXGAVLILITLFTLPGYK 840
Qy 841 TLLSRFLMWLCYLLTLGEAMVQEMAPMVQVGRGDGIIMAVAFYGVVFDITKMLAVL 900
Db 841 TLLSRFLMWLCYLLTLGEAMVQEMVPMQARGGRDGIIMAVATIFCGVVFDTIKMLAVL 900
Qy 901 GPAYLLKGLTRVPYFVRAHALLRMCTMAHLAGRRVOMALLALGRMTGTIYDHLTBM 960
Db 901 GPGLLRGLALTRVPYFVRAHALLRMCTVHHLAGRRVOMALLALGRMTGTIYDHLTBM 960
Qy 961 SDMAASGLRLAVAVEPIISPMCKKVIYVGAETACGDIHLGPVSARLGRGRIILGPAD 1020
Db 961 SDMAANGLRLAVAVEPIISPMCKKVIYVGAETACGDIHLGPVSARLGRGRIILGPAD 1020
Qy 1021 GYTSKGMSLAPITAYVQOTRGLGITIYVMTGRDKEQAGEIQVISTVQTSFLGTSIG 1080
Db 1021 GYTSKGMRLAPITAYVQOTRGLGIALVSMTRDKEQAGEIQVISTVQTSFLGTSIG 1080
Qy 1081 VLVMTVYHAGNKTLAGSRGPVTQWYSSAEGDLVGMPSPGTKSLBPTCGADVLYVTBN 1140
Db 1081 VLVMTVYHAGNKTLAGSRGPVTQWYSSAEGDLVGMPSPGTKSLBPTCGADVLYVTBN 1140
Qy 1141 ADVIPARRBGDKRGALLSPRLSTLKSSGGPVLCPRGHAGVGFRAAVCSRGVAKSIDPI 1200
Db 1141 ADVIPARRBGDKRGALLSPRLSTLKSSGGPVLCPRGHAGVGFRAAVCSRGVAKSIDPI 1200
Qy 1201 PVETLIDIVTRSPTFSDNSTPRAVPQTYQVGYLHAPFGSGSKTVPVAYAAQGYKVLVNP 1260
Db 1201 PVETLIDIVTRSPTFSDNSTPRAVPQTYQVGYLHAPFGSGSKTVPVAYAAQGYKVLVNP 1260
Qy 1261 SVAATLGFAGVYLSKAGINENIRTVARTVTGAPITVYTYGKFLADGGCAGAYDIIICD 1320
Db 1261 SVAATLGFAGVYLSKAGINENIRTVARTVTGAPITVYTYGKFLADGGCAGAYDIIICD 1320
Qy 1321 ECHAVDSTTLTGCTVUDQAFETAGVRLTVLATAIPRGSVTTBHPNIEEVALGQEGEIPRY 1380
Db 1321 ECHAVDSTTLTGCTVUDQAFETAGVRLTVLATAIPRGSVTTBHPNIEEVALGQEGEIPRY 1380
Qy 1381 GRAIPLSYIKGGRHLIFCHSKKKCDELAALRGKGLNSVAYYRGDLVSVIPTQGDVVVA 1440
Db 1381 GRAIPLSYIKGGRHLIFCHSKKKCDELAALRGKGLNSVAYYRGDLVSVIPTQGDVVVA 1440
Qy 1441 TDALMTGYTGDPRSVIDCNVAVATQVVDPSLDPTFTTTQIVPODAVSRSQRGRTRGRU 1500
Db 1441 TDALMTGYTGDPRSVIDCNVAVATQVVDPSLDPTFTTTQIVPODAVSRSQRGRTRGRU 1500
Qy 1501 GIIRVYSTGERASGMPDSVVLCECYDAGAMVELTSETTVRLRAVFNTPGLFVQCDDHE 1560
Db 1501 GIIRVYSTGERASGMPDSVVLCECYDAGAMVELTSETTVRLRAVFNTPGLFVQCDDHE 1560
Qy 1561 FMEAVFTGLTHIDAHFLSQTROSGENFAVLTAYQATVCARAKAPPSMDVMWKCILRLKP 1620
Db 1561 FMEAVFTGLTHIDAHFLSQTROSGENFAVLTAYQATVCARAKAPPSMDVMWKCILRLKP 1620
Qy 1621 TLVGPFTPLVRLGSVTNEVTULHPVTKYIATCQADLEWMTSTWVLAGGVLAAYVYCLA 1680
Db 1621 TLVGPFTPLVRLGSVTNEVTULHPVTKYIATCQADLEWMTSTWVLAGGVLAAYVYCLA 1680
Qy 1681 TGCVCITIGRLHINORAVVADKVELYAPDEMBECASRAALIEGGRILEMLESKIQGLL 1740
Db 1681 TGCVCITIGRLHINORAVVADKVELYAPDEMBECASRAALIEGGRILEMLESKIQGLL 1740
Qy 1741 QQASKQADIOPTVOASMFKVEQFWAKGMNFIISGLOYLAGLSTLPGNPAVASMMAFSA 1800
Db 1741 QQASKQADIOPTVOASMFKVEQFWAKGMNFIISGLOYLAGLSTLPGNPAVASMMAFSA 1800
Qy 1801 LTPPLSTSTIILNIIIGGLASQIAPAPAGATGVVSGLVGAAGVSGIGLKVLDVILLAGYG 1860
Db 1801 LTPPLSTSTIILNIIIGGLASQIAPAPAGATGVVSGLVGAAGVSGIGLKVLDVILLAGYG 1860
Qy 1861 AGISGALVAFKIMGKEMEDVNNLIGLISPGALVVGIVCAIILRRHVGPBEGAVQNM 1920
Db 1861 AGISGALVAFKIMGKEMEDVNNLIGLISPGALVVGIVCAIILRRHVGPBEGAVQNM 1920

Qy 1921 NRIIFASRGNHVAPTHVTESDASQRTYQLLGSITLTLRLRHLNMTTEDCPIPCGGSW 1960
Db 1921 NRIIFASRGNHVAPTHVTESDASQRTYQLLGSITLTLRLRHLNMTTEDCPIPCAGSW 1960
Qy 1961 LRDVDMVCTIITLDFKMLTSKLPKXMPGLPVSQCKQYKQVWAGTGMTTRCGANIS 2040
Db 1961 LRDVDMVCTIITLDFKMLTSKLPKXMPGLPVSQCKQYKQVWAGTGMTTRCGANIS 2040
Qy 2041 GNVRLGSMRITGPCKCMIMOGTFRINCYTGQCCPKRPAKPAKVAIMVVASEVAYTQH 2100
Db 2041 GNVRLGSMRITGPCKCMIMOGTFRINCYTGQCCPKRPAKPAKVAIMVVASEVAYTQH 2100
Qy 2101 GSYYITGLTITDNLKVPQCLPSPEFFSVWDQOIRHAPTRKPPFRDEVSFCVGLNSFV 2160
Db 2101 GSYYITGLTITDNLKVPQCLPSPEFFSVWDQOIRHAPTRKPPFRDEVSFCVGLNSFV 2160
Qy 2161 GSQLPCEPPTDVLMSMLTDPSSHITAETAAARLARGSPSEASASASQLSAPSLRATCT 2220
Db 2161 GSQLPCEPPTDVLMSMLTDPSSHITAETAAARLARGSPSEASASASQLSAPSLRATCT 2220
Qy 2221 THGRADVDMVDANLFWGCVYTRIESGKVYVLDLDPMBEERSDLEPSISEYMLPKR 2280
Db 2221 THGRADVDMVDANLFWGCVYTRIESGKVYVLDLDPMBEERSDLEPSISEYMLPKR 2280
Qy 2281 FPPALPAMARPDYNPVLVESMKRPDYCPATVAGCALPPRRKTPTPPPRRRTVGLSBDSI 2340
Db 2281 FPPALPAMARPDYNPVLVESMKRPDYCPATVAGCALPPRRKTPTPPPRRRTVGLSBDSI 2340
Qy 2341 GDALQOLAIKSRGPPSGDSGLSTGAGADSGQTPPEDELALSETGSISSMPLEGLG 2400
Db 2341 ADALQOLAIKTFGQPPSGDSGLSTGAGADSGQTPPEDELALSETGSISSMPLEGLG 2400
Qy 2401 DPELPEQVEQPPQGGVAAVPGSDSGSMSTCSEBDSDVCCSMSYMGALITPCSEB 2460
Db 2401 DPELPEQVEQPPQGGVAAVPGSDSGSMSTCSEBDSDVCCSMSYMGALITPCSEB 2460
Qy 2461 EKLPIINPLNSILRYHNKVYCTTTKSASLRAKKVTFDMQVLDVSYDVLNDIKLAASKV 2520
Db 2461 EKLPIINPLNSILRYHNKVYCTTTKSASLRAKKVTFDMQVLDVSYDVLNDIKLAASKV 2520
Qy 2521 TABLLTMEBAQCLTPPHSARSKYGFKAKEVNSLSGRAVNHIKSVKDLLEBSEPIPTTI 2580
Db 2521 TABLLTMEBAQCLTPPHSARSKYGFKAKEVNSLSGRAVNHIKSVKDLLEBSEPIPTTI 2580
Qy 2581 MAKNEVFCVDPPTKGGKKAARLIVPDLGVRCEKALVDIOKLPQAVMGASYGFOJSPA 2640
Db 2581 MAKNEVFCVDPPTKGGKKAARLIVPDLGVRCEKALVDIOKLPQAVMGASYGFOJSPA 2640
Qy 2641 QVBEFLKAMAEKKDPMGFSYDTRCFDSTVTERDITREESTYRACSLPEBAHTAISHLTE 2700
Db 2641 QVBEFLKAMAEKKDPMGFSYDTRCFDSTVTERDITREESTYRACSLPEBAHTAISHLTE 2700
Qy 2701 RLYVGGPMFNKSGQTCGRRCRASGVLTTSNGNTTCYVKLAACKAAGIAPMYVCGD 2760
Db 2701 RLYVGGPMFNKSGQTCGRRCRASGVLTTSNGNTTCYVKLAACKAAGIAPMYVCGD 2760
Qy 2761 DLVVISSESGTBEEDRNLRAPTEAMTRYSAPGDPREYVLELITSCSSVSVSVALGPQG 2820
Db 2761 DLVVISSESGTBEEDRNLRAPTEAMTRYSAPGDPREYVLELITSCSSVSVSVALGPQG 2820
Qy 2821 RRRVYLTBDPTTPIARAAMEVTRHSVNSWLGNIIOYAPTIWAMVMTLTHFFSILMAODT 2880
Db 2821 RRRVYLTBDPTTPIARAAMEVTRHSVNSWLGNIIOYAPTIWAMVMTLTHFFSILMAODT 2880
Qy 2881 LDONLNFEMYGAVSVSLDLPALIERLJGDASGLTYTHELTRVASALRKIGAPLR 2940
Db 2881 LDONLNFEMYGAVSVSLDLPALIERLJGDASGLTYTHELTRVASALRKIGAPLR 2940
Qy 2941 AMKSRARAVRASLSRGGRAVCGRYLFNMVVKTKLXLPBARI.LDLSMFTVGAAGG 3000
Db 2941 AMKSRARAVRASLSRGGRAVCGRYLFNMVVKTKLXLPBARI.LDLSMFTVGAAGG 3000

Db 1020 GYTSKGMRLIAPITAYAQOTRGLLGTI VVSMTRGDKTEQAGEI QVLTSTVTOSEFGTSLAG 1079
Qy 1081 VLTVTYHAGNKTLAGSGPVTOMYSAAEGDLVGMPSPPCTKSLJEBCTCGAVDLXLVTEN 1140
Db 1080 VLTMTVHAGNKTLAGSGPVTOMYSAAEGDLVGMSPGCTKSLJEBCTCGAVDLXLVTEN 1139
Qy 1141 ADVI PARRRDKRGALLSPRLSTLKGSSGGPVLCPRGHAGVGFRAAVCSRGVAKSIDPI 1200
Db 1140 ADVI PARRRDKRGALLSPRLSTLKGSSGGPVLCPRGHAGVGFRAAVCSRGVAKSIDPI 1199
Qy 1201 PVETLIDIVRSPTFESONSSTPEAVPOTYOVGVYLAHPGSGSKSTVPVPAVYAAGKVLVLP 1260
Db 1200 PVETLIDIVRSPTFESONSSTPEAVPOTYOVGVYLAHPGSGSKSTVPVPAVYAAGKVLVLP 1259
Qy 1261 SVAAITLGFAYLSKAGINENIRTVGRTVTGAPITVSTYKFLADGGAGAGAYDIIICD 1320
Db 1260 SVAAITLGFAYLSKAGINENIRTVGRTVTGAPITVSTYKFLADGGAGAGAYDIIICD 1319
Qy 1321 ECHAVDSTTLTIGIYVLDQAEITAGVRLTVLATITPPGSVTTTPHNTIEVALGGEIPIFY 1380
Db 1320 ECHAVDSTTLTIGIYVLDQAEITAGVRLTVLATITPPGSVTTTPHNTIEVALGGEIPIFY 1379
Qy 1381 GRAIPLSTYIKGGRHLIFCHSKKKCDELAALRGMGINSVAYYRGDLVSVIPTQGDVVVVA 1440
Db 1380 GRAIPLSTYIKGGRHLIFCHSKKKCDELAALRGMGINSVAYYRGDLVSVIPTQGDVVVVA 1439
Qy 1441 TDALMTGYTDPDSVIDCNVAVTQVDFSLDPTFTITTOIVPQDVAVSQRGRTRGRHL 1500
Db 1440 TDALMTGYTDPDSVIDCNVAVTQVDFSLDPTFTITTOIVPQDVAVSQRGRTRGRHL 1499
Qy 1501 GIYRYVSTGERASGMPDSVVLCECYDGAAMVYELTSETTVRLRAYFNTPGLPVCDHLE 1560
Db 1500 GIYRYVSTGERASGMPDSVVLCECYDGAAMVYELTSETTVRLRAYFNTPGLPVCDHLE 1559
Qy 1561 FMEAVETGLTHIDAHFLSOTKSGENPARYLTAYOATVCAARAKPPSPMDVMKCLTRLRP 1620
Db 1560 FMEAVETGLTHIDAHFLSOTKSGENPARYLTAYOATVCAARAKPPSPMDVMKCLTRLRP 1619
Qy 1621 TLVGPPTLYRLGSLVTNEVTLTHPVTKYIATCMQADLEVMTSTWVLAVGVLAAVAYCLA 1680
Db 1620 TLVGPPTLYRLGSLVTNEVTLTHPVTKYIATCMQADLEVMTSTWVLAVGVLAAVAYCLA 1679
Qy 1681 TGCVCCTIIGRLHINQRAVAVAPDKSVLYBAPDEMEECASRAALIEGGRIAEMLSKIQGLL 1740
Db 1680 TGCVCCTIIGRLHINQRAVAVAPDKSVLYBAPDEMEECASRAALIEGGRIAEMLSKIQGLL 1739
Qy 1741 QOASKAODIOPVVOASWPKVBOFWAKHMMNFISGIOYLAGSTLPGNPAVASMMAFSA 1800
Db 1740 QOASKAODIOPVVOASWPKVBOFWAKHMMNFISGIOYLAGSTLPGNPAVASMMAFSA 1799
Qy 1801 LTPSLSTSTTLINILIGMILASQIAPPAGATGFVBSGLVGAAGVSIQLGKVLVDILAAGY 1860
Db 1800 LTPSLSTSTTLINILIGMILASQIAPPAGATGFVBSGLVGAAGVSIQLGKVLVDILAAGY 1859
Qy 1861 AGISGALVAFKINSGEKPEMEDVYNLLPGILSPGALVGVYCAIIRRHVGEGBEAGVQNM 1920
Db 1860 AGISGALVAFKINSGEKPEMEDVYNLLPGILSPGALVGVYCAIIRRHVGEGBEAGVQNM 1919
Qy 1921 NRIIAFASRGNHVAPTHYVESDASORVTOGLSLITSLRLHNMITEDECI PCGSGM 1980
Db 1920 NRIIAFASRGNHVAPTHYVESDASORVTOGLSLITSLRLHNMITEDECI PCGSGM 1979
Qy 1981 LRDVMDVAVCTILDPFKNMLTSKLPFKMPLGLPVSCKGKGVAGVAGTOIMTTRPCGANIS 2040
Db 1980 LRDVMDVAVCTILDPFKNMLTSKLPFKMPLGLPVSCKGKGVAGVAGTOIMTTRPCGANIS 2039
Qy 2041 GNVRLGSMRITGPKTCMNIQGTFPINCYTEGOCVPEKAPNFEVVAIMRWAASEYAEVTOH 2100
Db 2040 GNVRLGSMRITGPKTCMNIQGTFPINCYTEGOCVPEKAPNFEVVAIMRWAASEYAEVTOH 2099
Qy 2101 GSYHYITGLTDLKVPCCOLPSPEFSPWVGVOIHRPAPRKPPFDEVSFCVGLNSFPV 2160
Db 2100 GSYHYITGLTDLKVPCCOLPSPEFSPWVGVOIHRPAPRKPPFDEVSFCVGLNSFPV 2159

Qy 2161 GSQLPCEDEPDDVLMWLTDPSSHITAEATARRLARGSPPEASASSASQLSAPSLRATCT 2220
Db 2160 GSQLPCEDEPDDVLMWLTDPSSHITAEATARRLARGSPPEASASSASQLSAPSLRATCT 2219
Qy 2221 THGKAYDVDMVDANLFMGGDVTRIBSGSKVVVLDSLDPMVEERDLEPSISEYMLPKR 2280
Db 2220 THGKAYDVDMVDANLFMGGDVTRIBSGSKVVVLDSLDPMVEERDLEPSISEYMLPKR 2279
Qy 2281 FPPALPAARPDVNPVLVESMKRPDYOPATVAGCALPPRRKTPPPPPRRRTVGSSEDSI 2340
Db 2280 FPPALPAARPDVNPVLVESMKRPDYOPATVAGCALPPRRKTPPPPPRRRTVGSSEDSI 2339
Qy 2341 GDALQOLAIKSFQGPPEPSGDSGLSTGACAADSGSQTPDEALALSETGISISSMPLJEGELG 2400
Db 2340 GDALQOLAIKSFQGPPEPSGDSGLSTGACAADSGSQTPDEALALSETGISISSMPLJEGELG 2399
Qy 2401 DDLPEPEQVEQPPPOGGVAAAPGSDSGSWSTCSEEDSVCCSNYSMTGALLITPCSE 2460
Db 2400 DDLPEPEQVEQPPPOGGVAAAPGSDSGSWSTCSEEDSVCCSNYSMTGALLITPCSE 2459
Qy 2461 EKLPIPNLSNLRHYNKVYCTTTKSASLRAKKVTFDPMQVLDYSYDSVLKDIKLAASKV 2520
Db 2460 EKLPIPNLSNLRHYNKVYCTTTKSASLRAKKVTFDPMQVLDYSYDSVLKDIKLAASKV 2519
Qy 2521 TARRLTMBEACQLTTPPHSARSKYGFAGKEVSLSGRAVNHIKSVWKDILLSEDETPPIPTI 2580
Db 2520 TARRLTMBEACQLTTPPHSARSKYGFAGKEVSLSGRAVNHIKSVWKDILLSEDETPPIPTI 2579
Qy 2581 MAKNEVFCVDPITKGGKKAARLIIVPDLCVRCERKALYDITQKLPQAVMGASVGFQVSPA 2640
Db 2580 MAKNEVFCVDPITKGGKKAARLIIVPDLCVRCERKALYDITQKLPQAVMGASVGFQVSPA 2639
Qy 2641 ORVEFLKAMAEKDPWGFSDYTRCFDSTVTERDITRTEESLYRACSLPEBAHTAIIHSLTE 2700
Db 2640 ORVEFLKAMAEKDPWGFSDYTRCFDSTVTERDITRTEESLYRACSLPEBAHTAIIHSLTE 2699
Qy 2701 RLUVGGMFNENSGOTCGYRRCRASGVLTTSMGNTITCYVKAALAAKKAAGIAPTMLVCGD 2760
Db 2700 RLUVGGMFNENSGOTCGYRRCRASGVLTTSMGNTITCYVKAALAAKKAAGIAPTMLVCGD 2759
Qy 2761 DLVVTSESGTBEDEERNIRATEAMTRYSAAPGPPREYDELEITSCSSNVSAVALGQ 2820
Db 2760 DLVVTSESGTBEDEERNIRATEAMTRYSAAPGPPREYDELEITSCSSNVSAVALGQ 2819
Qy 2821 RRRYVLTDPPTPIARAAMETVRHSVPVSMGNTIIQVAPTIWAKVWLTTHFFSILMAODT 2880
Db 2820 RRRYVLTDPPTPIARAAMETVRHSVPVSMGNTIIQVAPTIWAKVWLTTHFFSILMAODT 2879
Qy 2881 LDONINFEYGAIVSVSPDLDPALIERLHGLDAESLHTYTPHELTRVVASALRKLGAPLR 2940
Db 2880 LDONINFEYGAIVSVSPDLDPALIERLHGLDAESLHTYTPHELTRVVASALRKLGAPLR 2939
Qy 2941 AMKSRARAVRASLSIRGGRAAVCGRYLFENAVKTKLKLTPPEARLLDLSMFTVAGAGG 3000
Db 2940 AMKSRARAVRASLSIRGGRAAVCGRYLFENAVKTKLKLTPPEARLLDLSMFTVAGAGG 2999
Qy 3001 DIYHSVSARAPRLILFGLLLFVGVGLFLPAR 3033
Db 3000 DIYHSVSARAPRLILILFLFVGVGLFLPAR 3032

RESULT 11
0912A3
ID 0912A3 PRELIMINARY; PRT, 3033 AA.
AC 0912A3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.

OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MDa-5.
RA Itakura U., Nagayama K., Enomoto N., Kurosaki M., Watanabe H.,
Saco C.,
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
RM EMBL: AF238484; AAF59943.1; -
RR PIR: S35631, S35631.
DR HSPD, Q8UTS1, ICWX.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0019028; C: viral capsid; IEA.
DR GO: GO:0019031; C: viral envelope; IEA.
DR GO: GO:0005524; F: ATP binding; IEA.
DR GO: GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO: GO:0003723; F: RNA binding; IEA.
DR GO: GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO: GO:0008236; F: serine-type peptidase activity; IEA.
DR GO: GO:0005198; F: structural molecule activity; IEA.
DR GO: GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO: GO:0006350; P: transcription; IEA.
DR GO: GO:0019079; P: viral genome replication; IEA.
DR GO: GO:0019087; P: viral transformation; IEA.
DR InterPro: IPR00345; Cytochrome_heme_BS.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR01545; DEAD/DEAH_N.
DR InterPro: IPR002522; HCV_capsid.
DR InterPro: IPR002521; HCV_core.
DR InterPro: IPR002519; HCV_env.
DR InterPro: IPR002531; HCV_NS1.
DR InterPro: IPR000745; HCV_NS4a.
DR InterPro: IPR001490; HCV_NS4b.
DR InterPro: IPR002868; HCV_NS5a.
DR InterPro: IPR002166; HCV_RDRP.
DR InterPro: IPR004109; Peptidase_S29.
DR InterPro: IPR009003; Pept_Ser_Cys.
DR InterPro: IPR002518; Pept_U39_HCV_NS2.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVit.
DR Pfam: PF01543; HCV_capsid; 1.
DR Pfam: PF01542; HCV_core; 1.
DR Pfam: PF01539; HCV_env; 1.
DR Pfam: PF01560; HCV_NS1; 1.
DR Pfam: PF01538; HCV_NS2; 1.
DR Pfam: PF02907; HCV_NS3; 1.
DR Pfam: PF01006; HCV_NS4a; 1.
DR Pfam: PF01001; HCV_NS4b; 1.
DR Pfam: PF01506; HCV_NS5a; 1.
DR Pfam: PF00998; Viral_RDRP; 1.
DR SMART: SM00487; DEXDC; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; Transmembrane.
SQ SEQUENCE 3033 AA; 329516 MW; E01DDE175644593 CRC64;
Query Match 95.0%; Score 15322; DB 2; Length 3033;
Best local Similarity 94.2%; Pred. No. 0;
Matches 2858; Conservative 79; Mismatches 96; Indels 0; Gaps 0;

Db 181 LLSCTTPEVSAAEVKNISTGYMTNDCTNDSITWOLQAAVLHVPCCVPECKYGNASQCM 240
QY PVSPPVAVORRGALTOGRTTHIDMVVMSATCSALTYGDLGGVWLAOMFVSPQHMF 300
Db 241 PVSPPVAVORHGLAQGRTHIDMVVMSATCSALTYGDLGGVWLAOMFVSPQHMF 300
QY VQDCNCSYIPGTTIGHRAMDMNMMNSPTATMILAYARVBEVILIDISGAMGVMGFLA 360
Db 301 VQDCNCSYIPGTTIGHRAMDMNMMNSPTATMILAYARVBEVILIDISGAMGVMGFLA 360
QY YFSMOGAAKAVVAILLLAAGVADATHTVSGSAOQTGRLTSLFDMPKPKQIOLVNTNSG 420
Db 361 YFSMOGAAKAVVAILLLAAGVADATHTVSGSAOQTGRLTSLFDMPKPKQIOLVNTNSG 420
QY HINFTALNCNDSLHNGFIASLPHYHSFSSSCPEMSACRSIEMRVQMGALQYEDNVTN 480
Db 421 HINFTALNCNDSLHNGFIASLPHYHSFSSSCPEMSACRSIEMRVQMGALQYEDNVTN 480
QY PEDMRPYCMHYPPRCQGVASAKTVCGPYCTPSPFVVVGTTRDLGAPTYTNGENETDVL 540
Db 481 PEDMRPYCMHYPPRCQGVASAKTVCGPYCTPSPFVVVGTTRDLGAPTYTNGENETDVL 540
QY PEDMRPYCMHYPPRCQGVASAKTVCGPYCTPSPFVVVGTTRDLGAPTYTNGENETDVL 540
Db 481 PEDMRPYCMHYPPRCQGVASAKTVCGPYCTPSPFVVVGTTRDLGAPTYTNGENETDVL 540
QY LNSTRPVIGSWPGCTWNNSSGYTTKCGAPPCRTADFNASTDLCPTPCFRKHPTTYLK 600
Db 541 LNSTRPVIGSWPGCTWNNSSGYTTKCGAPPCRTADFNASTDLCPTPCFRKHPTTYLK 600
QY LNSTRPVIGSWPGCTWNNSSGYTTKCGAPPCRTADFNASTDLCPTPCFRKHPTTYLK 600
Db 541 LNSTRPVIGSWPGCTWNNSSGYTTKCGAPPCRTADFNASTDLCPTPCFRKHPTTYLK 600
QY CGSGFWLTPKCLVDVYLYMHVPCVTNTIKIRIVYGVVHRLTAACNPTRGDCNLED 660
Db 601 CGSGFWLTPKCLVDVYLYMHVPCVTNTIKIRIVYGVVHRLTAACNPTRGDCNLED 660
QY RDRSOLSPILHSTTEMALIPCSYSDLPALSTGLLHONIVDVOFMYGSPALTKYIVRM 720
Db 661 RDRSOLSPILHSTTEMALIPCSYSDLPALSTGLLHONIVDVOFMYGSPALTKYIVRM 720
QY RDRSOLSPILHSTTEMALIPCSYSDLPALSTGLLHONIVDVOFMYGSPALTKYIVRM 720
Db 661 RDRSOLSPILHSTTEMALIPCSYSDLPALSTGLLHONIVDVOFMYGSPALTKYIVRM 720
QY EWMVILLFLTLADARVACLMWILLGQEAALBKLVILHAASASCSNGFLYFVFPVAM 780
Db 721 EWMVILLFLTLADARVACLMWILLGQEAALBKLVILHAASASCSNGFLYFVFPVAM 780
QY EWMVILLFLTLADARVACLMWILLGQEAALBKLVILHAASASCSNGFLYFVFPVAM 780
Db 721 EWMVILLFLTLADARVACLMWILLGQEAALBKLVILHAASASCSNGFLYFVFPVAM 780
QY YIKGRVAVPLATYSLTGLMSPFLILLALPOQAVYADASVHGOIGALLVMTLFTLTPYK 840
Db 781 YIKGRVAVPLATYSLTGLMSPFLILLALPOQAVYADASVHGOIGALLVMTLFTLTPYK 840
QY YIKGRVAVPLATYSLTGLMSPFLILLALPOQAVYADASVHGOIGALLVMTLFTLTPYK 840
Db 781 YIKGRVAVPLATYSLTGLMSPFLILLALPOQAVYADASVHGOIGALLVMTLFTLTPYK 840
QY TLLSRFLMWLCYLLTLGEMTQEMAPPMQVGRGDIIMAVAFYPGVVFDTIKMLAVL 900
Db 841 TLLSRFLMWLCYLLTLGEMTQEMAPPMQVGRGDIIMAVAFYPGVVFDTIKMLAVL 900
QY TLLSRFLMWLCYLLTLGEMTQEMAPPMQVGRGDIIMAVAFYPGVVFDTIKMLAVL 900
Db 841 TLLSRFLMWLCYLLTLGEMTQEMAPPMQVGRGDIIMAVAFYPGVVFDTIKMLAVL 900
QY GPAYLLKGAULTRPVYFVRAHALLRMCNTHALAGRVYQOMALLGRWTGYIYDHLTPM 960
Db 901 GPAYLLKGAULTRPVYFVRAHALLRMCNTHALAGRVYQOMALLGRWTGYIYDHLTPM 960
QY GPAYLLKGAULTRPVYFVRAHALLRMCNTHALAGRVYQOMALLGRWTGYIYDHLTPM 960
Db 901 GPAYLLKGAULTRPVYFVRAHALLRMCNTHALAGRVYQOMALLGRWTGYIYDHLTPM 960
QY SDMAASGLRDLAVAEPIIFSPMEKRVIVMGAETPACGDIILHGLPVSARLREVLGGPAD 1020
Db 961 SDMAASGLRDLAVAEPIIFSPMEKRVIVMGAETPACGDIILHGLPVSARLREVLGGPAD 1020
QY SDMAASGLRDLAVAEPIIFSPMEKRVIVMGAETPACGDIILHGLPVSARLREVLGGPAD 1020
Db 961 SDMAASGLRDLAVAEPIIFSPMEKRVIVMGAETPACGDIILHGLPVSARLREVLGGPAD 1020
QY GYTSKMGSLAPITRAYAOQTGRLGAIYVSMGTGRDKTEQAGEIQLSTVYTSFGLTSSG 1080
Db 1021 GYTSKMGSLAPITRAYAOQTGRLGAIYVSMGTGRDKTEQAGEIQLSTVYTSFGLTSSG 1080
QY GYTSKMGSLAPITRAYAOQTGRLGAIYVSMGTGRDKTEQAGEIQLSTVYTSFGLTSSG 1080
Db 1021 GYTSKMGSLAPITRAYAOQTGRLGAIYVSMGTGRDKTEQAGEIQLSTVYTSFGLTSSG 1080
QY VLMVTYHAGNKTLTAGSRGPTOMYSAGDLVGMSPDPGTSLEPCTCGADVLLVTRN 1140
Db 1081 VLMVTYHAGNKTLTAGSRGPTOMYSAGDLVGMSPDPGTSLEPCTCGADVLLVTRN 1140
QY VLMVTYHAGNKTLTAGSRGPTOMYSAGDLVGMSPDPGTSLEPCTCGADVLLVTRN 1140
Db 1081 VLMVTYHAGNKTLTAGSRGPTOMYSAGDLVGMSPDPGTSLEPCTCGADVLLVTRN 1140
QY ADVIPARRGDKRGLSPRLSTLKSGSGGPVCPRHAAVCPVPAAYCSRVAASIDPI 1200
Db 1141 ADVIPARRGDKRGLSPRLSTLKSGSGGPVCPRHAAVCPVPAAYCSRVAASIDPI 1200
QY ADVIPARRGDKRGLSPRLSTLKSGSGGPVCPRHAAVCPVPAAYCSRVAASIDPI 1200
Db 1141 ADVIPARRGDKRGLSPRLSTLKSGSGGPVCPRHAAVCPVPAAYCSRVAASIDPI 1200
QY FVETLDIYTRSPTESDNTSPPAVPOTYOVGYLAHPTSGSKTKVPVAAAGSYKVLVLP 1260
Db 1201 FVETLDIYTRSPTESDNTSPPAVPOTYOVGYLAHPTSGSKTKVPVAAAGSYKVLVLP 1260
QY FVETLDIYTRSPTESDNTSPPAVPOTYOVGYLAHPTSGSKTKVPVAAAGSYKVLVLP 1260
Db 1201 FVETLDIYTRSPTESDNTSPPAVPOTYOVGYLAHPTSGSKTKVPVAAAGSYKVLVLP 1260
QY SVAATLGFAGVAYSKAGINPNIIRGTGRTVTGAPITTSYFGKPLADGGCAGADYIICD 1320
Db 1261 SVAATLGFAGVAYSKAGINPNIIRGTGRTVTGAPITTSYFGKPLADGGCAGADYIICD 1320
QY SVAATLGFAGVAYSKAGINPNIIRGTGRTVTGAPITTSYFGKPLADGGCAGADYIICD 1320
Db 1261 SVAATLGFAGVAYSKAGINPNIIRGTGRTVTGAPITTSYFGKPLADGGCAGADYIICD 1320

```

Qy 1321 ECHAVDSITILIGIVLDOAETAGVRLTVLATATPPGVSVTTTPHPIIEEVALGOBEIEPIY 1380
Db 1321 ECHAVDATITLIGIVLDOAETAGVRLTVLATATPPGVSVTTTPHPIIEEVALGOBEIEPIY 1380
Qy 1381 GBAIPLSYIKGGRHLIFCHSKKKCEBLAALRGKGLNSVAYRRGDVSIIPFGSVVYVA 1440
Db 1381 GBAIPLSYIKGGRHLIFCHSKKKCEBLAALRGKGLNSVAYRRGDVSIIPFGSVVYVA 1440
Qy 1441 TALALMTGYTDPDSVIDCNVAATVQVVDPSLDPFTITTOIVPODAVSRSQRGRTGRGL 1500
Db 1441 TALALMTGYTDPDSVIDCNVAATVQVVDPSLDPFTITTOIVPODAVSRSQRGRTGRGL 1500
Qy 1501 GIYRVYSTGERASGMEDSVLCECYDAGAAWYELTPESETTVLRAYFNTPGD.PVCODHLE 1560
Db 1501 GIYRVYSTGERASGMEDSVLCECYDAGAAWYELTPESETTVLRAYFNTPGD.PVCODHLE 1560
Qy 1561 FHEAVFTGLTHIDAHFLSOTKSGENFALITAYQATVCARAKAPPSMVMKCLTRLP 1620
Db 1561 FHEAVFTGLTHIDAHFLSOTKSGENFALITAYQATVCARAKAPPSMVMKCLTRLP 1620
Qy 1621 TLVGPFTLLYRLGVSVNEVTLTHPVTKYIATCMQADLEWMTSTWVLGGVLAAYACLA 1680
Db 1621 TLVGPFTLLYRLGVSVNEVTLTHPVTKYIATCMQADLEWMTSTWVLGGVLAAYACLA 1680
Qy 1681 TSCVCTIIGRLHINQRAVVAADKEVLYEAFDEMEBCASRAALIIEGQRIAEMLKSKIQGL 1740
Db 1681 TSCVCTIIGRLHINQRAVVAADKEVLYEAFDEMEBCASRAALIIEGQRIAEMLKSKIQGL 1740
Qy 1741 QOASKAQAOIOPVQASWPKRQ/PQAKHMMNFI.SGI.QYI.AGSLTLP.GNNAVVA.SMA.FSA 1800
Db 1741 QOASKAQAOIOPVQASWPKRQ/PQAKHMMNFI.SGI.PYLARLSQ.PGNNAVVA.SMA.FKA 1800
Qy 1801 LPSPLSTSTIILNIGMLASQIAPAGATGVVSGVGAAGVSGIGLKVVDIAGY 1860
Db 1801 LPSPLSTSTIILNIGMLASQIAPAGATGVVSGVGAAGVSGIGLKVVDIAGY 1860
Qy 1861 AGISGALVAFKIMSGEKPSMEDVNNLLPGLSPGALVGVICAILRRHVGEAGVQMM 1920
Db 1861 AGISGALVAFKIMSGEKPSMEDVNNLLPGLSPGALVGVICAILRRHVGEAGVQMM 1920
Qy 1921 NRIIAASRGNHVAPTHYTESDASQRYQLGSLITSLRLRHHMTIEDPPIPCAGSW 1980
Db 1921 NRIIAASRGNHVAPTHYTESDASQRYQLGSLITSLRLRHHMTIEDPPIPCAGSW 1980
Qy 1981 LRDVMDVCTIILDFKQMLTSLKLPKMPGL.PFVSCQKGYGVAWGIMTTRCPGCANIS 2040
Db 1981 LRDVMDVCTIILDFKQMLTSLKLPKMPGL.PFVSCQKGYGVAWGIMTTRCPGCANIS 2040
Qy 2041 GNVRLGSMRITGPKTCMNIWQTFPINCYTEGQCVBPAPNFKVALMRVAASEVAVTQH 2100
Db 2041 GNVRLGSMRITGPKTCMNIWQTFPINCYTEGQCVBPAPNFKVALMRVAASEVAVTQH 2100
Qy 2101 GSVHYITGLTNDLKVPCQLPSEPFSSWDGVOIHRFAATPKPFFDEVSFCVGLNSFV 2160
Db 2101 GSVHYITGLTNDLKVPCQLPSEPFSSWDGVOIHRFAATPKPFFDEVSFCVGLNSFV 2160
Qy 2161 GSOLPCDPEPDTVLSMMLTDPESHITAETAAARLARGSPSEASASSASQLSAPSLRATCT 2220
Db 2161 GSOLPCDPEPDTVLSMMLTDPESHITAETAAARLARGSPSEASASSASQLSAPSLRATCT 2220
Qy 2221 THGKAYDVMDVNDALFMGADVTRIESGSKVVLDSIDPWEERSDLEPSIPSEYMLPKKR 2280
Db 2221 THGKAYDVMDVNDALFMGADVTRIESGSKVVLDSIDPWEERSDLEPSIPSEYMLPKKR 2280
Qy 2281 FPPALPAMARPDVNPPLVBSMKRPDQOPATVACALPPRRKPTTPPPRRRRRTVGLSEDSI 2340
Db 2281 FPPALPAMARPDVNPPLVBSMKRPDQOPATVACALPPRRKPTTPPPRRRRRTVGLSEDSI 2340
Qy 2341 GVALQOLAIKSFQCPSPGDSGLSTGAGAADSQOTPPPELALSERGSISSMPPLEGGT 2400
Db 2341 GVALQOLAIKSFQCPSPGDSGLSTGAGAADSQOTPPPELALSERGSISSMPPLEGGT 2400

```

```

Qy 2401 DPLLEPQVEPQPPPGGVAAFGSDSGSWTCSSEEDDSVCCSNYSYMTGALITPCSPBE 2460
Db 2401 DPLLEPQVEPQPPPGGVAAFGSDSGSWTCSSEEDDSVCCSNYSYMTGALITPCSPBE 2460
Qy 2461 EKLPIPLNSLILYHNKVVCTTTSKASLRAKKTTPRMQVLDVSYOSVUKDITLAASKV 2520
Db 2461 EKLPIPLNSLILYHNKVVCTTTSKASLRAKKTTPRMQVLDVSYOSVUKDITLAASKV 2520
Qy 2521 TABLLTMEACOLTPPHSASKYFGKAYEVSLSGRAVNHISYVAKLLEDESEPIPTTI 2580
Db 2521 NARLLTYEACOLTPPHSASKYFGKAYEVSLSGRAVNHISYVAKLLEDESEPIPTTI 2580
Qy 2581 MAKNEVFCVDPFKGKKAARLIYVPI.DGVRCERMAIYDITOKL.PQAVMGASYSFOYSPA 2640
Db 2581 MAKNEVFCVDPFKGKKAARLIYVPI.DGVRCERMAIYDITOKL.PQAVMGASYSFOYSPA 2640
Qy 2641 QRVFELLKAMAEKDDPMGFSDTFCFSTVTERDIRTEESIYQACSLPEEARTAIHSLTE 2700
Db 2641 QRVFELLKAMAEKDDPMGFSDTFCFSTVTERDIRTEESIYQACSLPEEARTAIHSLTE 2700
Qy 2701 RLIVYGMFNSKQTCGYRRCRAGSVLITSMGNTITCYVKALAAKKAAGIAPTMLVCGD 2760
Db 2701 RLIVYGMFNSKQTCGYRRCRAGSVLITSMGNTITCYVKALAAKKAAGIAPTMLVCGD 2760
Qy 2761 DLVVISSQGTEDERNIRAFTEAMTRYSAPEPDPPEYDLELITSCSNVSVVALGPOG 2820
Db 2761 DLVVISSQGTEDERNIRAFTEAMTRYSAPEPDPPEYDLELITSCSNVSVVALGPOG 2820
Qy 2821 RRRYYLFRDPTTPARAAWETVRHSPVNSW.GNI.IQVAPITWAMVLMTHFEFSLMODT 2880
Db 2821 RRRYYLFRDPTTPARAAWETVRHSPVNSW.GNI.IQVAPITWAMVLMTHFEFSLMODT 2880
Qy 2881 LDQNLNFEYMGAVSVSPLDPAIIERLHGLDASLHTYTPHELTRVVASALRKLGAPELR 2940
Db 2881 LDQNLNFEYMGAVSVSPLDPAIIERLHGLDASLHTYTPHELTRVVASALRKLGAPELR 2940
Qy 2941 AMKSRARAARASLISRGRAAVCGRYLFNMAVTKLXITLPEARLLDLSSWFTVGAAGG 3000
Db 2941 AMKSRARAARASLISRGRAAVCGRYLFNMAVTKLXITLPEARLLDLSSWFTVGAAGG 3000
Qy 3001 DIYHSVSRARPRLLIFGLLLFVGVGLFLPAR 3033
Db 3001 DIYHSVSRARPRLLIFGLLLFVGVGLFLPAR 3033

RESULT 12
Q912A2 PRELIMINARY; PRT; 3033 AA.
AC Q912A2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Polypeptin.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepatitis.
OX NCBI_Taxid=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MDa-7;
RA Itakura J., Nagayama K., Enomoto N., Kurosaki M., Watanabe H.,
RA Sato C.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF238485; AAF59944.1; -.
DR HSSP; 08JYSL; 1CMX;
DR GO; GO:001602; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.

```

DR GO:0005198; F:structural molecule activity; IEA.
 DR GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO:0006530; P:transcription; IEA.
 DR GO:0019079; P:viral genome replication; IEA.
 DR GO:0019087; P:viral transformation; IEA.
 DR InterPro: IPR000345; Cyt_c_heme_BS.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR011545; DEAD/DEAH_N.
 DR InterPro: IPR002522; HCV_capsid.
 DR InterPro: IPR002521; HCV_core.
 DR InterPro: IPR002519; HCV_env.
 DR InterPro: IPR002511; HCV_NSI.
 DR InterPro: IPR000745; HCV_NS4a.
 DR InterPro: IPR001490; HCV_NS4b.
 DR InterPro: IPR002668; HCV_NS5a.
 DR InterPro: IPR002166; HCV_RdRp.
 DR InterPro: IPR004109; Peptidase_S29.
 DR InterPro: IPR009003; Pept_Ser_Cys.
 DR InterPro: IPR002518; Pept_U39_HCV_NS2.
 DR InterPro: IPR007095; RNA_pol_DS_PS.
 DR InterPro: IPR007094; RNA_pol_Psivir.
 DR Pfam: PF01543; HCV_core; 1.
 DR Pfam: PF01542; HCV_core; 1.
 DR Pfam: PF01539; HCV_env; 1.
 DR Pfam: PF01538; HCV_NSI; 1.
 DR Pfam: PF01538; HCV_NSI; 1.
 DR Pfam: PF02907; HCV_NS3; 1.
 DR Pfam: PF01006; HCV_NS4a; 1.
 DR Pfam: PF01001; HCV_NS4b; 1.
 DR Pfam: PF01506; HCV_NS5a; 1.
 DR Pfam: PF00998; Viral_RdRp; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR PROSITE: PS00190; CYTOCHROME C; UNKNOWN 1.
 DR Coar protein; Envelope protein; Glycoprotein; Nonstructural protein.
 DR PolyProtein; Transmembrane.
 KW POLYPEPTIDE; 3033 AA; 329532 MW; BE135578FABA1E87 CRC64;
 SQ SEQUENCE

Query Match 95.0%; Score 15321; DB 2; Length 3033;
 Beat Local Similarity 94.0%; Pred. No. 0;
 Matches 2852; Conservative 88; Mismatches 93; Indels 0; Gaps 0;

QY 1 MSTNKPORKTKRNTNRNRPODVKFPQGGQIVGGVLLPRGRPLGVRATRKTSERSQPRG 60
 DB 1 MSTNKPORKTKRNTNRNRPODVKFPQGGQIVGGVLLPRGRPLGVRATRKTSERSQPRG 60
 QY 61 RRQPIPKRRRSTGKSGKRGYWPPLYGNBGLCMAGWLLSPGSRSPSGWGNDRHRSRNVG 120
 DB 61 RRQPIPKRRRSTGKSGKRGYWPPLYGNBGLCMAGWLLSPGSRSPSGWGNDRHRSRNVG 120
 QY 121 KYIDLTGCGFADLMGYIPVVGAPLGAVARALAHGVRVLEDDGVNPFATGNLPGCSFSIFILA 180
 DB 121 KYIDLTGCGFADLMGYIPVVGAPLGAVARALAHGVRVLEDDGVNPFATGNLPGCSFSIFILA 180
 QY 181 LLSCITTEVSAAEVKNISTGVYVNDCTNDSITWQLQAALVHVPICVPCSEKVGASQCVI 240
 DB 181 LLSCITTEVSAAEVKNISTGVYVNDCTNDSITWQLQAALVHVPICVPCSEKVGASQCVI 240
 QY 241 PVSERVAVQRBALTOGLRTHIDMYVMSATLCSALYVGDLCGGVWLAAMFIVSQHMF 300
 DB 241 PVSERVAVQRBALTOGLRTHIDMYVMSATLCSALYVGDLCGGVWLAAMFIVSQHMF 300
 QY 301 VQECNCSTYLGITTHRRAMDMMMMSPTTMTILAVMRVPEVITDIISGAMGVMFGIA 360
 DB 301 VQECNCSTYLGITTHRRAMDMMMMSPTTMTILAVMRVPEVITDIISGAMGVMFGIA 360
 QY 361 YFSOGAAKVVVILLLAGVDARTHTVGGSAAGTTGRLTSLFDMGPROKIQLVNTNSG 420
 DB 361 YFSOGAAKVVVILLLAGVDARTHTVGGSAAGTTGRLTSLFDMGPROKIQLVNTNSG 420
 QY 421 HINRFLALNCNSLHGTPLASLFTYHSFNSSGCEPMSACRSTEARVCGALQYEDNTN 480
 DB 421 HINRFLALNCNSLHGTPLASLFTYHSFNSSGCEPMSACRSTEARVCGALQYEDNTN 480
 QY 480 HINRFLALNCNSLHGTPLASLFTYHSFNSSGCEPMSACRSTEARVCGALQYEDNTN 480
 DB 480 HINRFLALNCNSLHGTPLASLFTYHSFNSSGCEPMSACRSTEARVCGALQYEDNTN 480

QY 481 PEDMRPCWHPYPRQCGVVSATKVGSPVYCTTPSPVYVGTTRDLGAPTYTGWENETDVL 540
 DB 481 PEDMRPCWHPYPRQCGVVSATKVGSPVYCTTPSPVYVGTTRDLGAPTYTGWENETDVL 540
 QY 541 LNSTRPPLGSMGCTWMSNGYTKTCGAPPCGTRADFNASTDULCPTDCFRKHPDTTYLK 600
 DB 541 LNSTRPPLGSMGCTWMSNGYTKTCGAPPCGTRADFNASTDULCPTDCFRKHPDTTYLK 600
 QY 601 CGSGPMLPRCLIDVPRYLMHYPCYVNYTIFKIRMYGVGEHRLTAACNFTGDRCNLMD 660
 DB 601 CGSGPMLPRCLIDVPRYLMHYPCYVNYTIFKIRMYGVGEHRLTAACNFTGDRCNLMD 660
 QY 661 RDRSQSPPLHSTTEMALIPCSYSDLPALSTGLHLHONIIVDQFMYGSPALTKYIVM 720
 DB 661 RDRSQSPPLHSTTEMALIPCSYSDLPALSTGLHLHONIIVDQFMYGSPALTKYIVM 720
 QY 721 EHVILLFLILDARVACIWMILLIGQEALEKLVILHAASACNGFLYFVYFVAA 780
 DB 721 EHVILLFLILDARVACIWMILLIGQEALEKLVILHAASACNGFLYFVYFVAA 780
 QY 781 YIKGRVPLATYSLTGLMFPCLLLALPOAAVAVASVQGLGVSILVLTFTLTPAPK 840
 DB 781 YIKGRVPLATYSLTGLMFPCLLLALPOAAVAVASVQGLGVSILVLTFTLTPAPK 840
 QY 841 TLLSRFLMWLCYLLTIGEMVQEMAPPMQVRGGRDGIWAVALPYPGVFDITKMLAVL 900
 DB 841 TLLSRFLMWLCYLLTIGEMVQEMAPPMQVRGGRDGIWAVALPYPGVFDITKMLAVL 900
 QY 901 GPAYLLKGLATRVYFVRALHLLRMCTMARHLAAGRYVOMALLAGRWGTYYIYDHLTM 960
 DB 901 GPAYLLKGLATRVYFVRALHLLRMCTMARHLAAGRYVOMALLAGRWGTYYIYDHLTM 960
 QY 961 SDMAASGLBDIAVAEPIIFSPMEKKVYVMGETAACGDIHLGLVSARLAGEVILLGPD 1020
 DB 961 SDMAASGLBDIAVAEPIIFSPMEKKVYVMGETAACGDIHLGLVSARLAGEVILLGPD 1020
 QY 1021 GYTSKQMSLAPITAYAOOTRGLGTIVSMGRDTEQAGEIQVLSYTVQSFGLTSISG 1080
 DB 1021 GYTSKQMSLAPITAYAOOTRGLGTIVSMGRDTEQAGEIQVLSYTVQSFGLTSISG 1080
 QY 1081 VLMTVYHGAANKTLTAGSRQPVTOYSSAEGDVGMPSPGTSLPSCGADVILVTRN 1140
 DB 1081 VLMTVYHGAANKTLTAGSRQPVTOYSSAEGDVGMPSPGTSLPSCGADVILVTRN 1140
 QY 1141 ADVIPARRRGDGRGALLSPRLSTLKSSGGGVLCPRGHAVGVFPAACSRVAASIDPT 1200
 DB 1141 ADVIPARRRGDGRGALLSPRLSTLKSSGGGVLCPRGHAVGVFPAACSRVAASIDPT 1200
 QY 1201 PVETLDIVTRSPPTFSDNSTPPAVPOTYOVYLAHPTGSGKSTKVVAAYAGYKVLVLP 1260
 DB 1201 PVETLDIVTRSPPTFSDNSTPPAVPOTYOVYLAHPTGSGKSTKVVAAYAGYKVLVLP 1260
 QY 1261 SVAATLGFAYLSKAGINPNIRTVGTGAPITYSYTGKFLADGGCAGAYDIICD 1320
 DB 1261 SVAATLGFAYLSKAGINPNIRTVGTGAPITYSYTGKFLADGGCAGAYDIICD 1320
 QY 1321 ECHAVDSTTIIGIGVLDQAEAGVRLTVLATPGRSVTPPHNIEEVALGOBESEIPY 1380
 DB 1321 ECHAVDSTTIIGIGVLDQAEAGVRLTVLATPGRSVTPPHNIEEVALGOBESEIPY 1380
 QY 1381 GRAIPLSYIKGGRHLIFGSKKKKCELAALRGMLNSVAAYRGADVSIPTQGVVVVA 1440
 DB 1381 GRAIPLSYIKGGRHLIFGSKKKKCELAALRGMLNSVAAYRGADVSIPTQGVVVVA 1440
 QY 1441 TDALMTGYTGDFSDYIDCNVAVTVVDFSLDPTFTTITQIYVPODVSRSORRGRTGRGL 1500
 DB 1441 TDALMTGYTGDFSDYIDCNVAVTVVDFSLDPTFTTITQIYVPODVSRSORRGRTGRGL 1500
 QY 1501 GIYRVYSTGERASGDFSVLCECYDAGAAWYELLPSTTYRLRAYFNTPGLPVQODHLE 1560
 DB 1501 GIYRVYSTGERASGDFSVLCECYDAGAAWYELLPSTTYRLRAYFNTPGLPVQODHLE 1560
 QY 1561 FWEAVFTGLTHIDAHFLSQTKQSGENFAYLTAQATVCARAKAPPSPDWVMKCLTRLKP 1620
 DB 1561 FWEAVFTGLTHIDAHFLSQTKQSGENFAYLTAQATVCARAKAPPSPDWVMKCLTRLKP 1620

Db 1561 FWEAVFGLTHIDAHFISQTKOGENFAYLVAAQATVCARAKAPPSWVMWMCILTRLRX 1620
Qy 1621 TLVGPPLRLRLSNTNEVTLTHPYKYIATCMQADLEWMTSWVLAVGVLAAVAYCXA 1680
Db 1621 TLVGPPLRLRLSNTNEVTLTHPYKYIATCMQADLEWMTSWVLAVGVLAAVAYCXA 1680
Qy 1681 TGCVCIIIGRLHINQRAVAVADKEVLYEAPDEMECASRAALIIEGORIAEMLKSKIQGLL 1740
Db 1681 TGCVCIIIGRLHINQRAVAVADKEVLYEAPDEMECASRAALIIEGORIAEMLKSKIQGLL 1740
Qy 1741 QOASKAQADIQPAVEASWPVEQFMAKMMNFISGIQYLAGLSTLIGNPAVASMAFSA 1800
Db 1741 QOASKAQADIQPAVEASWPVEQFMAKMMNFISGIQYLAGLSTLIGNPAVASMAFSA 1800
Qy 1801 LTSPLSTSTILNIIIGMILASQIAPAGATGVVSGLVGAAGVSGIGLVGLVDILAGYG 1860
Db 1801 LTSPLSTSTILNIIIGMILASQIAPAGATGVVSGLVGAAGVSGIGLVGLVDILAGYG 1860
Qy 1861 AGISGALVAFKIMSGEKSPMEDVYNLLPGILSPGALVGVICAGILRRHVGPGEAGVQW 1920
Db 1861 AGISGALVAFKIMSGEKSPMEDVYNLLPGILSPGALVGVICAGILRRHVGPGEAGVQW 1920
Qy 1921 NRLLAFASRGNHVAPTHYTESDAQRVTLQSLITSLRLRLHMTEDCFIPGAGSM 1980
Db 1921 NRLLAFASRGNHVAPTHYTESDAQRVTLQSLITSLRLRLHMTEDCFIPGAGSM 1980
Qy 1981 LRQVMDVAVCTIILDFKMWLTSKLFPKMPGLPFVSCOKGYVAGVAGIMTRPCGANIS 2040
Db 1981 LRQVMDVAVCTIILDFKMWLTSKLFPKMPGLPFVSCOKGYVAGVAGIMTRPCGANIS 2040
Qy 2041 GNVRLGSMRITGPKTCNNINQGTPIPCYTEGQCVCPKPAFKAIAIRVAASEYAEVTH 2100
Db 2041 GNVRLGSMRITGPKTCNNINQGTPIPCYTEGQCVCPKPAFKAIAIRVAASEYAEVTH 2100
Qy 2101 GSHYITGLTTDNLKVPQCLPSPFESWVDVQIHRPAPPKPPFDEVSFCGLNSFVY 2160
Db 2101 GSHYITGLTTDNLKVPQCLPSPFESWVDVQIHRPAPPKPPFDEVSFCGLNSFVY 2160
Qy 2161 GSQLPCEPBDPTVLMSMLTDPHSITETAARLARLSPSESSASSASQLSAPSLATCT 2220
Db 2161 GSQLPCEPBDPTVLMSMLTDPHSITETAARLARLSPSESSASSASQLSAPSLATCT 2220
Qy 2221 THGKAYDVMDVANDLFWGSDVTRIESSGKVVVLDLDPWEERSDLEPSIPSEYMLPKKR 2280
Db 2221 THGKAYDVMDVANDLFWGSDVTRIESSGKVVVLDLDPWEERSDLEPSIPSEYMLPKKR 2280
Qy 2281 FPPALPAMAPPDYNNPPLVESWKPDPYOPATVAGCALPPPKKTYTPPPRRRTYGLSEDT 2340
Db 2281 FPPALPAMAPPDYNNPPLVESWKPDPYOPATVAGCALPPPKKTYTPPPRRRTYGLSEDT 2340
Qy 2341 GDALQOLAIKSPFGPPSGSGSGSTGAGADSGSQPPDLALSETGSSMPPLGEG 2400
Db 2341 GDALQOLAIKSPFGPPSGSGSGSTGAGADSGSQPPDLALSETGSSMPPLGEG 2400
Qy 2401 DDPLEBPQVPRPPPGGVAAPSGDSGWSCTSEBDSVCCSMYSWTCALITPCSPRE 2460
Db 2401 DDPLEBPQVPRPPPGGVAAPSGDSGWSCTSEBDSVCCSMYSWTCALITPCSPRE 2460
Qy 2461 EKPPIPLNSLRLRYHNKVCCTTKASLBAKKYTPRMQVLDSDYSDVUKIKLAASKY 2520
Db 2461 EKPPIPLNSLRLRYHNKVCCTTKASLBAKKYTPRMQVLDSDYSDVUKIKLAASKY 2520
Qy 2521 TARLLTWEAAGCLTPHSARSKYGFKAVERSLSGRVNHIKSWKDLLEDSETPPTT 2580
Db 2521 SARLLTWEAAGCLTPHSARSKYGFKAVERSLSGRVNHIKSWKDLLEDSETPPTT 2580
Qy 2581 MAKNVEVCVDPPTGKGKKAARLIYVDPDGVAVCEKMAIYDTOKLPAVWAGSAGFOYSPA 2640
Db 2581 MAKNVEVCVDPPTGKGKKAARLIYVDPDGVAVCEKMAIYDTOKLPAVWAGSAGFOYSPA 2640
Qy 2641 QRVFELLKAAEKKDPMGFSYDTRCFDSTVTERDIRTEBIIYACSLPEEAHAIHSLTE 2700
Db 2641 QRVFELLKAAEKKDPMGFSYDTRCFDSTVTERDIRTEBIIYACSLPEEAHAIHSLTE 2700

Db 2641 QRVFELLKAAEKKDPMGFSYDTRCFDSTVTERDIRTEBIIYACSLPEEAHAIHSLTE 2700
Qy 2701 RLYVGGPMFNSKQTCGRRCASGVLTTSMGNITTCYVKAACAKAGIAPTMLVCGD 2760
Db 2701 RLYVGGPMFNSKQTCGRRCASGVLTTSMGNITTCYVKAACAKAGIAPTMLVCGD 2760
Qy 2761 DLVVISSEGTEDERNLRAFTAMTRYSAAPGDPREYDLELTICSSNVSVAGPQG 2820
Db 2761 DLVVISSEGTEDERNLRAFTAMTRYSAAPGDPREYDLELTICSSNVSVAGPQG 2820
Qy 2821 RRRYLTDDPTTPRAAMETVRSHPVNSWGNIIQVAPTIWVAVMTTFESTIMADOT 2880
Db 2821 RRRYLTDDPTTPRAAMETVRSHPVNSWGNIIQVAPTIWVAVMTTFESTIMADOT 2880
Qy 2881 LDQDINFEMGCVSVSPDLPAIIEHAGDAFLHHTYTHELTRVASALRKLGAPELR 2940
Db 2881 LDQDINFEMGCVSVSPDLPAIIEHAGDAFLHHTYTHELTRVASALRKLGAPELR 2940
Qy 2941 AKKSPARAARASLSIRGGRAAVCGRYLFNMAVKTKLTPLEPEARLILSSWFTVAGCGG 3000
Db 2941 AKKSPARAARASLSIRGGRAAVCGRYLFNMAVKTKLTPLEPEARLILSSWFTVAGCGG 3000
Qy 3001 DIYHSVRARPRLLFGILLFVGGLFLPAR 3033
Db 3001 DIYHSVRARPRLLFGILLFVGGLFLPAR 3033
RESULT 13
Q99B0 PRELIMINARY; PRT; 3033 AA.
AC Q99B0;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_Taxid=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:21361470; PubMed=11468731, DOI=10.1002/jmv.1073;
RA Kurthara C., Ishiyama N., Nishiyama Y., Fukushi S., Kageyama T.,
RT "Molecular characterization of hepatitis C virus genotype 2a from the
RT entire sequences of four isolates."
RL J. Med. Virol. 64:466-475(2001).
DR EMBL: AF163002; AAF2610.1; -.
DR HSP; OBYS1; ICWX.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; F:transcription; IEA.
DR GO; GO:0019079; F:viral genome replication; IEA.
DR GO; GO:0019087; F:viral transformation; IEA.
DR InterPro: IPR000345; Cyclic_heme_BS.
DR InterPro: IPR001410; DEAD_BS.
DR InterPro: IPR001545; DEAD/DEAH_N.
DR InterPro: IPR002522; HCV_core.
DR InterPro: IPR002521; HCV_core.
DR InterPro: IPR002519; HCV env.
DR InterPro: IPR002531; HCV NS1.
DR InterPro: IPR000745; HCV NS4a.
DR InterPro: IPR001490; HCV NS4b.
DR InterPro: IPR002868; HCV NS5a.
DR InterPro: IPR002166; HCV_RdRP.

DR InterPro, IPR004109, Peptidase_S29.
DR InterPro, IPR009003, Pept_Ser_Cys.
DR InterPro, IPR002518, Pept_U39_HCV_NS2.
DR InterPro, IPR007095, RNA_pol_DS_PS.
DR InterPro, IPR007094, RNA_pol_PsVtrL.
DR Pfam, PF01543, HCV_capsid, 1.
DR Pfam, PF01542, HCV_core, 1.
DR Pfam, PF01539, HCV_env, 1.
DR Pfam, PF01560, HCV_NS1, 1.
DR Pfam, PF01538, HCV_NS2, 1.
DR Pfam, PF02907, HCV_NS3, 1.
DR Pfam, PF01006, HCV_NS4, 1.
DR Pfam, PF01001, HCV_NS4b, 1.
DR Pfam, PF01506, HCV_NS5a, 1.
DR Pfam, PF00998, Viral_RdRp, 1.
DR SMART, SMO0487, DEXDC, 1.
DR PROSITE, PS00190, CYTOCHROME_C, UNKNOWN_1.
KW Coat protein, Envelope protein, Glycoprotein, Nonstructural protein,
KW Polyprotein, Transmembrane.
QO POLYSEQUENCE 3033 AA; 329180 MW; 6E3F51FD3ABA4C24 CRC64;

Query Match	95.0%;	Score 15311;	DB 2;	Length 3033;
Best Local Similarity	94.1%;	Pred. No. 0;		
Matches 2855;	Conservative 75;	Mismatches 103;	Indels 0;	Gaps 0;

Qy	1	MSINPKPORKTKRNTNRRPQDVKPFQCGOIIVCGVYLIPRGRPLGVARTKTSSENSQPRG	60
Dp	1	MSINPKPORKTKRNTNRRPODVKPFQCGOIVCGVYLIPRGRPLGVARTKTSSENSQPRR	60
Qy	61	RROP1PKORRSTGSKMGKPGVMP1LYGNGLGMAGLILS PRSRSRPMKGNDRPHRRRNNG	120
Dp	61	RROP1PKORRSTGRSGKRGV1WPLVXNGGLGMAGLILS PRSRSRPMKGNDRPHRRRNNG	120
Qy	121	KVIDTLTCGFADLMGY1PVVGAPLGCVABALAHGVRVIEDGVNPFATGNLPGCSFS1FLLA	180
Dp	121	KVIDTLTCGFADLMGY1PVVGAPLGCVATLALAHGVRVLEDEGINFATGNLPGCSFS1FLLA	180
Qy	181	LLSCITTVSAAEVENISTGVWVNTDCTNDS1TMOLOAAVLHVPQGVPEKXGNASQCMT	240
Dp	181	LLSCITTVSAAQOVNNTSGVWVNTDNCADS1TMOLOAAVLHVPQGVPEKXXYTSRCWI	240
Qy	241	PVSPVNAVQRPQALTOGLRTH1DMVYMSATLTCALYVGDLCGGVMAAQMFIYSPQHMF	300
Dp	241	PVSPVNAVROGALTOGLRTH1DMVYMSATLTCALYVGDLCGGVMAAQMFIYSPBHMF	300
Qy	301	VQDCNCSTYRPGITIGHRAAMDMMNMSPRATMILAYANRVEPV1DI1ISGAHGVNFGLA	360
Dp	301	VOECNCSTYPGAITQORAMADMMNMSPRATMILAYANRVEPV1DI1ISGAHGVNFGLA	360
Qy	361	YFSMGGAAKAVVILLILAAGVDARTHTVCGSAAQTGRLTSLFVDMGPROKIOLVNTNGSM	420
Dp	361	YFSMGGAAKAVVILLILAAGVDANTXXXGGAARNAAYGTSLSFXGAXON1QOLINTNGSM	420
Qy	421	HINRTALNCDSLATGFTASLFYTHS FNSSGCPDERMSACRS1EAPRVGMALOYEDVNTN	480
Dp	421	HINRTALNCDSLNTNGTFLAXLFYXNRFNSSGCPERLVCRN1EAFR1IGWTLQOYEDNVFN	480
Qy	481	PEDMRPYCMH1PPRCGCVGSAKTVCGPVYCFPTSPVVGVTYTRLGAP1TYTNGENEDVFL	540
Dp	481	PEDMRPYCMH1PPKCSITVPARSVCGPVYCFPTSPVVGVTYTRBRGV1PTYTWGENEDVFL	540
Qy	541	LNSTRBPLGSMFGCTWANSRGYTKTCGAPPCSTRADFNASTDULCPTDCPRKHDPDTTYLK	600
Dp	541	LNSTRBPGSGWFGCTWANSRGYTKTCGAPPCSTRADFNASTDULCPTDCFRKHPEATYLK	600
Qy	601	CGSGGWLTPRCCLIDVPYRLMHYPC1VNVY1FXIRMYVGVGERLTAACNFTGDRCNLED	660
Dp	601	CGSGGWLTPRCCLIDVPYRLMHYPC1VNVY1FXIRMYVGVGERLTAACNFTGDRCDLED	660
Qy	661	RDRSGLSPLLHSTHTEMAILPCSYSDLPALSTGLHLHONI1VDVQEMYGSLPALTKY1VM	720
Dp	661	RDRSGLSPLLHSTHTEMAILPCSYSDLPALSTGLHLHONI1VDVQVYAGSLPALTKY1VM	720

QY	721	EWVILLFLLLDADAVOCIMMLILGQEALELKVLIHAASASCMGFLYFVIFPAAM	780
Db	721	EMVVL:FLLLDADAVOCIMMFIIMGQEALEKLVLIHAASASCMGFLYFVIFPAAM	780
QY	781	YIKRVPLATYSITGMSFSLLLALPQOAYADASHGOIGAAALVMI:LFLLTGYK	840
Db	781	YIKRPAVPLAAYSITGMSFCLLLVLPQOAYAEASVHGOIGAAALVLI:LFLLTGYK	840
QY	841	TLSRFIMWL:CYLLTLGEAMQEWAPPQVGRGDIIMAWAIFYPGVFDITKMLAVL	900
Db	841	TLSRCLMWCYLLTLGEPAMQEWAPSMQARGDGIIMATITPCGVFDITKMLAAL	900
QY	901	GPAYILKALTRVPEFPRAAHLLMCTMARHLAGRYOAAALLAGRTGYIYDHLTPM	960
Db	901	GSGYLLR:GALTRVPEFPRAAHLLMCTIAKLAGKAYOAMALLAGRTGYIYDHLTPM	960
QY	961	SDMAASGRDIAVVEPII:PSPMKXYIWMGAETPAACGDILHG:PVASRIGREVLGPAD	1020
Db	961	SDMAASGRDIAVVEPII:PSPMKXYIWMGAETPAACGDILHG:PVASRIGREVLGPAD	1020
QY	1021	GYSKSGMSLAPITAYAOOTRGLGTTIVSMGTGRDKTEOAGELOVLSTVTSF:IGTSISG	1080
Db	1021	GYSKSGMRLLAPITAYAOOTRGLGALIVSMGTGRDKTEOAGELOVLSTVTSF:IGTSISG	1080
QY	1081	VLMTVYHAGNKTLAGSRGPTQWYSSAEGDLVGPSP:PGTKSL:EPCTCGAVDLYLVTRN	1140
Db	1081	VLMTVYHAGNKTLAGSRGPTQWYSSAEGDLVGPSP:PGTKSL:EPCTCGAVDLYLVTRN	1140
QY	1141	ADVI:PARRGKRAALLSPRL:STLKSSGGPV:CPRGHANGVRAVCSGVAKS:IDFI	1200
Db	1141	ADVI:PARRGKRAALLSPRL:STLKSSGGPV:CPRGHANGVRAVCSGVAKS:IDFI	1200
QY	1201	PVETLIDITRSP:TESDNSTPRPAVQOTQVGLIAP:PTSCKSTKPVAVYAAQGYKVLVLP	1260
Db	1201	PVETLIDITRSP:TESDNSTPRPAVQOTQVGLIAP:PTSCKSTKPVAVYASQYKVLVLP	1260
QY	1261	SVAATLFGAYLSRAHGINPNI:RTGVRTVTGADITY:STYCKFLADGCGAGAYDIIICD	1320
Db	1261	SVAATLFGAYLSRAHGINPNI:RTGVRTVTGAIATYSTYCKFLADGCGAGAYDIIICD	1320
QY	1321	ECHAVDSTTIIIGITVLDQAE:TAGVRLTVLATAP:PPSGSYTTPHNIIEVALGOGEL:PFY	1380
Db	1321	ECHAVDATTIIIGITVLDQAE:TAGVRLTVLATAP:PPSPSYTTPHNIIEVALGHEGEL:PFY	1380
QY	1381	GRAI:PLASYIKKGRHLIFCHSKKCCDELAAL:RGMLSVAYYRSLDYSV:PTQGDVVVVA	1440
Db	1381	GKA:PLSYIKKGRHLIFCHSKKCCDELSAAL:RGMLAANAAYIRKLDYSAL:PAQGDVVVVA	1440
QY	1441	TDALMTGYTGD:PDSDIVDCNAVTVQVDFSLDPTTITTOI:VPDQAVSRSQRGRGTGRGL	1500
Db	1441	TDALMTGYTGD:PDSDIVDCNAVTVQVDFSLDPTTITTOI:VPDQAVSRSQRGRGTGRGL	1500
QY	1501	GIYRYVSTGERASGMFDSVULCECYDGAAMYELT:SETTYRLRAYINTPGL:PCODHLE	1560
Db	1501	GIYRYVSTGERASGMFDSVULCECYDGAAMYELT:PAETTYRLRAYINTPGL:PCODHLE	1560
QY	1561	FWEAVFTGLTHIDAHFISOTKQSENAFYI:LAYOATYCARAKAP:PPSMDVMKCL:TRLKP	1620
Db	1561	FWEAVFTGLTHIDAHFISOTKQSENAFYI:LAYOATYCARAKAP:PPSMDVMKCL:TRLKP	1620
QY	1621	TLVGPPTLLYRLGSVTEVTLTHPVTKYIATCMQAD:LEVMTSTWVLAGVLA:AAVAA:CLA	1680
Db	1621	TLVGPPTLLYRLGSVTEVTLTHPVTKYIATCMQAD:LEVMTSTWVLAGVLA:AAVAA:CLA	1680
QY	1681	TGCYCIIIGRLHINQRAYAPKEXLYEAFDEMEC:ASRAALIIEGQRI:IAELKSKIOGLL	1740
Db	1681	TGCYSIIIGRLHINQRAYAPKEXLYEAFDEMEC:ASRAALIIEGQRI:IAELKSKIOGLL	1740
QY	1741	QOAKROADIODPTQOASMPKVEQWAGHMNFISIGI:OYLAGLSTLPCNPAPAYASMA:FSAA	1800
Db	1741	QOAKROADIODPAQASMPKVEQWAGHMNFISIGI:OYLAGLSTLPCNPAPAYASMA:FSAA	1800
QY	1801	LTPSLSTSTTILNITLGWLASQIAPAGATGPFVSGV:AGAAVSGSIGI:KVLVDIILAGYG	1860

[illegible]

Db	2881	LDQUNLEPMGCAVSVSPLDLPAIERLHGDAESLHTYTHELTRVAALRLKIGAPRLR	2944
Qy	2941	AMKSRARAVRASLLISRGRAAVCGRRYLFENNAVTKTLKLTPLPEARLLDLSMFTVGAGG	3000
Db	2941	AMKSRARAVRASLLISRGRAAVCGRRYLFENNAVTKTLKLTPLPEARHDLSSMFTVGAGG	3000
Qy	3001	DIHVSARARPRLLLFGLLLFVGVGLFLPAR	3033
Db	3001	DIHVSARARPRLLLFGLLLFVGVGLFLPAR	3033
RESULT 14			
Q9Q9A8			
AC	Q9Q9A8	PRELIMINARY;	PRT; 3033 AA.
DT	01-MAY-2000	(TRENBLrel. 13, Created)	
DT	01-MAY-2000	(TRENBLrel. 13, Last sequence update)	
DT	01-MAR-2004	(TRENBLrel. 26, Last annotation update)	
DE	Polyprotein.		
OS	Hepatitis C virus.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;		
OC	Hepadnaviruses.		
OX	NCBI_Taxid=11103;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21361470; PubMed=11466731; DOI=10.1002/jmv.1073;		
RA	Kurihara C., Ishiyama N., Nishiyama Y., Fukushima S., Kageyama T.,		
RA	Katayama K., Miura S.;		
RT	"Molecular characterization of hepatitis C virus genotype 2a from the		
RT	entire sequences of four isolates";		
RL	J. Med. Virol. 64:466-475(2001).		
DR	EMBL; AF169004; AAF25612.1; -.		
DR	HSSP; Q8JYS1; 1CMX.		
DR	GO: GO:0016021; C: integral to membrane; IEA.		
DR	GO: GO:0019028; C: viral capsid; IEA.		
DR	GO: GO:0019031; C: viral envelope; IEA.		
DR	GO: GO:0005524; F: ATP binding; IEA.		
DR	GO: GO:0008026; F: ATP-dependent helicase activity; IEA.		
DR	GO: GO:0003723; F: RNA binding; IEA.		
DR	GO: GO:0003968; F: RNA-directed RNA polymerase activity; IEA.		
DR	GO: GO:0008236; F: serine-type peptidase activity; IEA.		
DR	GO: GO:0005198; F: structural molecule activity; IEA.		
DR	GO: GO:0006508; F: proteolysis and peptidolysis; IEA.		
DR	GO: GO:0006350; P: transcription; IEA.		
DR	GO: GO:0019079; P: viral genome replication; IEA.		
DR	GO: GO:0019087; P: viral transformation; IEA.		
DR	InterPro; IPR000345; Cytc_hem_BS.		
DR	InterPro; IPR001410; DEAD.		
DR	InterPro; IPR011545; DEAD/DEAH_N.		
DR	InterPro; IPR002522; HCV capsid.		
DR	InterPro; IPR002521; HCV core.		
DR	InterPro; IPR002519; HCV env.		
DR	InterPro; IPR002531; HCV NS1.		
DR	InterPro; IPR000745; HCV NS4a.		
DR	InterPro; IPR001490; HCV NS4b.		
DR	InterPro; IPR002868; HCV NS5a.		
DR	InterPro; IPR002166; HCV RdRp.		
DR	InterPro; IPR004109; Peptidase_S29.		
DR	InterPro; IPR009003; Pept_Ser_Cys.		
DR	InterPro; IPR002518; Pept_U39_HCV_NS2.		
DR	InterPro; IPR007095; RNA_pol_DS_PS.		
DR	InterPro; IPR007094; RNA_pol_PSVir.		
DR	Pfam; PF01543; HCV_capsid; 1.		
DR	Pfam; PF01542; HCV_core; 1.		
DR	Pfam; PF01539; HCV_env; 1.		
DR	Pfam; PF01560; HCV_NS1; 1.		
DR	Pfam; PF01538; HCV_NS2; 1.		
DR	Pfam; PF02907; HCV_NS3; 1.		
DR	Pfam; PF01006; HCV_NS4a; 1.		
DR	Pfam; PF01001; HCV_NS4b; 1.		
DR	Pfam; PF01506; HCV_NS5a; 1.		
DR	Pfam; PF00998; Viral_RdRp; 1.		
DR	SMART; SMO0487; DEXDc; 1.		

DR PROSITE, PS00190; CYTOCHROME C; UNKNOWN 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KM Polyprotein; Transmembrane.
 SQ SEQUENCE 3033 AA; 329175 MW; 74932A91979E78A5 CRC64;

Query Match 94.7%; Score 15266; DB 2; Length 3033;
 Best Local Similarity 94.0%; Pred. No. 0;
 Matches 2850; Conservative 71; Mismatches 112; Indels 0; Gaps 0;

```

QY 1 MSTNPKPKTKRNTNRPRPODVKPPGGQIVGGVILLPRGPRLCVGRATRKTSERSQPRG 60
DB 1 MSTKRPQKTKRNTNRPRPODVKPPGGQIVGGVILLPRGPRLCVGRATRKTSERSQPRG 60
QY 61 RRQPIPKDRRSTGKSGWKGPMPPLXGNBGLMAGWLLSPGSRSPSMGPNDRHSRRNG 120
DB 61 RRQPIPKDRRSTGKSGWKGPMPPLXGNBGLMAGWLLSPGSRSPSMGPNDRHSRRNG 120
QY 121 KVIDTLTGGFADLMGYIPVVGAPLGGVAPALAHGVRLVEDGVNFATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTGGFADLMGYIPVVGAPLGGVAPALAHGVRLVEDGVNFATGNLPGCSFSIFLLA 180
QY 181 LLSCTTPPSAAEVNISTGVNVTNDCTNDSITWQLOAAVLHVPGCVPEKYGANSQCYI 240
DB 181 LLSCTTIPVSAQVKNSTGYNVTNDCSNDSTWQLOAAVLHVPGCVPEKYGANSRCWI 240
QY 241 PVSPPVAVQVRPGALTOGLRTHIDMTVMASATLCSALYVGDLCGVMILAAQMFIVSQHMF 300
DB 241 PVSPPVAVQVRPGALTOGLRTHIDMTVMASATLCSALYVGDLCGVMILAAQMFIVSQHMF 300
QY 301 VQDCNCSIYPGTITGHRMAMDMNMNNSPATWILAYAKRPEVIIIDICGALHMGVFGLA 360
DB 301 VQECNCSIYPGTITGHRMAMDMNMNNSPATWILAYAKRPEVIIIDICGALHMGVFGLA 360
QY 361 YFSMGAAKVVVILLLAGVDARTHTVGSAAQTGRLTSLFDMGPRKTOLVNTNGSW 420
DB 361 YFSMGAAKVVVILLLAGVDARTHTVGSAAQTGRLTSLFDMGPRKTOLVNTNGSW 420
QY 421 HINRTALNCNDSLHTGFIASLFTYTHFNSSGCPERRMSACRSIEAFRVGALQYEDNTN 480
DB 421 HINRTALNCNDSLHTGFIASLFTYTHFNSSGCPERRMSACRSIEAFRVGALQYEDNTN 480
QY 481 PEDMRPYCMHYPPROCGVNSAKTVCGPVYCTTPSPVNVGTITDRLGAPITYWGENETVEL 540
DB 481 PEDMRPYCMHYPPROCGVNSAKTVCGPVYCTTPSPVNVGTITDRLGAPITYWGENETVEL 540
QY 541 LNSTPRPLGSGWGCCTMMNNSGYTKTCGAPPCRTTRADFNASTDILLPTDCFRAGHPDTYIK 600
DB 541 LNSTPRPLGSGWGCCTMMNNSGYTKTCGAPPCRTTRADFNASTDILLPTDCFRAGHPDTYIK 600
QY 601 CGSGPMLTPRCLIDYPRILMHYPCVTNNTIFKIRMYVGVEHRLTAAACNFTGRGCNLED 660
DB 601 CGSGPMLTPRCLIDYPRILMHYPCVTNNTIFKIRMYVGVEHRLTAAACNFTGRGCNLED 660
QY 661 RDRSGLSPLHSTTEMAILPCSYDLPALSTGLHLHQNIVDQVMYGLSPALTRYIYRW 720
DB 661 RDRSGLSPLHSTTEMAILPCSYDLPALSTGLHLHQNIVDQVMYGLSPALTRYIYRW 720
QY 721 EMVILLFLILLDARVACCLIMLILGOAFAALEKIVLHAASASCNGLFVIVFPVAM 780
DB 721 EMVILLFLILLDARVACCLIMLILGOAFAALEKIVLHAASASCNGLFVIVFPVAM 780
QY 781 YIKGKXVPLATYSLTGLMFSLLALPQOAYAYASVHGOIGALLIILITLFTLPAYK 840
DB 781 YIKGKXVPLATYSLTGLMFSLLALPQOAYAYASVHGOIGALLIILITLFTLPAYK 840
QY 841 TLLSFLMWLCYLLTLGEMVQEMAPPMQVGRGDGIIMAVAI FYPGVFDITKMLAVL 900
DB 841 TLLSFLMWLCYLLTLGEMVQEMAPPMQVGRGDGIIMAVAI FYPGVFDITKMLAVL 900
QY 901 GPAYLLKALTPVYFVRAHALBMCMAHLAAGRYVOMALLAGRTGTITYLHLP 960
DB 901 GPAYLLKALTPVYFVRAHALBMCMAHLAAGRYVOMALLAGRTGTITYLHLP 960
QY 961 GGYLLRDLALTRVYFVRAHALBMCMAHLAAGRYVOMALLAGRTGTITYLHLP 960
DB 961 GGYLLRDLALTRVYFVRAHALBMCMAHLAAGRYVOMALLAGRTGTITYLHLP 960

```

```

QY 961 SDMAASGLRLAVALVEPIIFSPMEKKVIWGAETAAACDILHGLPVSARLIGREVLLGPD 1020
DB 961 SDMAASGLRLAVALVEPIIFSPMEKKVIWGAETAAACDILHGLPVSARLIGREVLLGPD 1020
QY 1021 GYTSKGMSLLAITAYAAQOTRGLLCTIVSMTGRDKTEOAGEIQLVSLVYQSFLGTSISG 1080
DB 1021 GYTSKGMSLLAITAYAAQOTRGLLCTIVSMTGRDKTEOAGEIQLVSLVYQSFLGTSISG 1080
QY 1081 VLMTVYHAGNKTLAGSRGPVLYOMYSABEGDLVGPSPGKSLSPCTCGAIDLVLN 1140
DB 1081 VLMTVYHAGNKTLAGSRGPVLYOMYSABEGDLVGPSPGKSLSPCTCGAIDLVLN 1140
QY 1141 ADVIPARRRGDKRGALLSPRPSTLKGSSGGFVLCPRGHAVGFRAAACSRSVAASIDFI 1200
DB 1141 ADVIPARRRGDKRGALLSPRPSTLKGSSGGFVLCPRGHAVGFRAAACSRSVAASIDFI 1200
QY 1201 PVEALAIYTRSPFTSDNSTPPAVPQTVGYLHAPTGSGKSTKVVAVAAQGYKVLVLNP 1260
DB 1201 PVEALAIYTRSPFTSDNSTPPAVPQTVGYLHAPTGSGKSTKVVAVAAQGYKVLVLNP 1260
QY 1261 SVAATLGGAYLSTKAGINPNIRTCGVRTVTGAPITYSTYGKFLADGGCAGAYDIICD 1320
DB 1261 SVAATLGGAYLSTKAGINPNIRTCGVRTVTGAPITYSTYGKFLADGGCAGAYDIICD 1320
QY 1321 ECHAVDSTTIIGIGIVLDOAETAGVRLTVLATATPPGCVTTPHPIIEEVALGOEIEIPY 1380
DB 1321 ECHAVDSTTIIGIGIVLDOAETAGVRLTVLATATPPGCVTTPHPIIEEVALGOEIEIPY 1380
QY 1381 GRAIPLSYIKGGRHLIFCHSKKKCDELAALRGMLNSVAYYRGLDVSVIPTQGDVVVA 1440
DB 1381 GRAIPLSYIKGGRHLIFCHSKKKCDELAALRGMLNSVAYYRGLDVSVIPTQGDVVVA 1440
QY 1441 TDALMTGYTGPDSYIDCNNAVTVQVYVDSLPFTTITQVPODVSSRQRRGRGRGL 1500
DB 1441 TDALMTGYTGPDSYIDCNNAVTVQVYVDSLPFTTITQVPODVSSRQRRGRGRGL 1500
QY 1501 GIYRYVSTGEBAASGMFDSVILCECYDAGAAMVELTPSESTYRLRAYFMTPGLPVCODHLE 1560
DB 1501 GIYRYVSTGEBAASGMFDSVILCECYDAGAAMVELTPSESTYRLRAYFMTPGLPVCODHLE 1560
QY 1561 FWEAVFTGLTHIDAFHLSQTKQSGENFAYLVAYQATVCARAKAPPSPWDIMKCLTRLPK 1620
DB 1561 FWEAVFTGLTHIDAFHLSQTKQSGENFAYLVAYQATVCARAKAPPSPWDIMKCLTRLPK 1620
QY 1621 TLVGPFTLLYRLGSVTNVTLTHPYTKYIAICOMQADLEVMSTWLAGVLAAVAACYLA 1680
DB 1621 TLVGPFTLLYRLGSVTNVTLTHPYTKYIAICOMQADLEVMSTWLAGVLAAVAACYLA 1680
QY 1681 TGCVCIIIGRLHINORAVVAPDKEVLYEAFDMEBCASRAALIIEGORIAEMLSKTIQGL 1740
DB 1681 TGCVCIIIGRLHINORAVVAPDKEVLYEAFDMEBCASRAALIIEGORIAEMLSKTIQGL 1740
QY 1741 QOASKOADIOPTVOASHPKYEOFWAKHMNPFISGIIQYLAGLSTLPGNPAAVAMAFSA 1800
DB 1741 QOASKOADIOPTVOASHPKYEOFWAKHMNPFISGIIQYLAGLSTLPGNPAAVAMAFSA 1800
QY 1801 LTPSLSTSTILLNLTGMLASQIAPPAGANGFVSGVGAAGVSGIGLXVLVDLIAAG 1860
DB 1801 LTPSLSTSTILLNLTGMLASQIAPPAGANGFVSGVGAAGVSGIGLXVLVDLIAAG 1860
QY 1861 AGISGALVAFKIMSGEKSPMEDVNLLEGLISPGALVGVCAALIRRHVPGEGAVOM 1920
DB 1861 AGISGALVAFKIMSGEKSPMEDVNLLEGLISPGALVGVCAALIRRHVPGEGAVOM 1920
QY 1921 NRIIAFASRGNHVAFTHVTESDASQRTQLLSITISLRLRLHMTTEDCPIPCAGSW 1980
DB 1921 NRIIAFASRGNHVAFTHVTESDASQRTQLLSITISLRLRLHMTTEDCPIPCAGSW 1980
QY 1981 LRDVMDVCSITLIDPKMNLSTKLPKMPGGLPFVSCQGYKQVMACTGIMTRCPCGANIS 2040
DB 1981 LRDVMDVCSITLIDPKMNLSTKLPKMPGGLPFVSCQGYKQVMACTGIMTRCPCGANIS 2040
QY 2041 GNVRLGSMRITGPCTCNMIWQSTFPPINCYTEGOCVPKPAFNFKVAIWRVAASEVAEVTQH 2100

```

```
|||||
Db 2041 GNRRLSMRITGPKTCMNTWGHGTFPINCYEGQCVKPA.BNFXTAIWRAVAASEYAEVTOH 2100
Qy 2101 GSVHYITGLTTNDLKYPCOLPSEPEFSWVGVOIHFAPRPKPFPEDEXFCGJNSFFV 2160
Db 2101 GSVHYITGLTTNDLKYPCOLPSEPEFSWVGVOIHFAPRPKPFPEDEXFCGJNSFFV 2160
Qy 2161 GSQLPCDPEEDTVLMSMLTDPESHITAETARLARLARGSPPESEASSASOLASPLRATCT 2220
Db 2161 GSQLPCDPEEDTVLMSMLTDPESHITAETARLARLARGSPPESEASSASOLASPLRATCT 2220
Qy 2221 THGKAVDVMDVNLFWGCDVTRIEGSKYVLDLSDPWVEERSDLEPSIPSEYMLPKKR 2280
Db 2221 THGKAVDVMDVNLFWGCDVTRIEGSKYVLDLSDPWVEERSDLEPSIPSEYMLPKKR 2280
Qy 2281 FPPALPAMARPDVNPPLVSESKRPDQAPVACALP.PPKPKTPPPRRRRRTGSGEDSI 2340
Db 2281 FPPALPAMARPDVNPPLVSESKRPDQAPVACALP.PPKPKTPPPRRRRRTGSGEDSI 2340
Qy 2341 GDALQOLAISKFGQPPPSGDSGLSTAGADSGSQTPPELALSETGSISSMPLEBELG 2400
Db 2341 GDALQOLAISKFGQPPPSGDSGLSTAGADSGSQTPPELALSETGSISSMPLEBELG 2400
Qy 2401 DPLPEBQVPEPQPPQGVAAPESSDSGSMSTCEEDSVYCCSMYSWTGALITPCSPBE 2460
Db 2401 DPLPEBQVPEPQPPQGVAAPESSDSGSMSTCEEDSVYCCSMYSWTGALITPCSPBE 2460
Qy 2461 EKLPIPLNSLRHYNNKYCTTTSKASLRAKAYTDRMQVLDSDYSDVLDKIDLAASY 2520
Db 2461 EKLPIPLNSLRHYNNKYCTTTSKASLRAKAYTDRMQVLDSDYSDVLDKIDLAASY 2520
Qy 2521 TABLLTMEBAQOLTPPHSARSKYFGAKKEYRSLSGRAVNHKISVWKDLEDSEPTPTT 2580
Db 2521 TABLLTMEBAQOLTPPHSARSKYFGAKKEYRSLSGRAVNHKISVWKDLEDSEPTPTT 2580
Qy 2581 MANGNEFCVDPTGSKKAARLIYPLDGVAVCEKMLYDTQCLPAVWAGSYGFOYSRA 2640
Db 2581 MANGNEFCVDPTGSKKAARLIYPLDGVAVCEKMLYDTQCLPAVWAGSYGFOYSRA 2640
Qy 2641 QRYEFLLKAAVEKKDPMGFSYDTRCFDSYTERDITTEESIYACSLPEBAKTAHSLTE 2700
Db 2641 QRYEFLLKAAVEKKDPMGFSYDTRCFDSYTERDITTEESIYACSLPEBAKTAHSLTE 2700
Qy 2701 RLTVGSGPMFNSKQTCGYRCRASGVLTSMGNTITCYVKAALACKAGIAPTMLVCGD 2760
Db 2701 RLTVGSGPMFNSKQTCGYRCRASGVLTSMGNTITCYVKAALACKAGIAPTMLVCGD 2760
Qy 2761 DLVVISSESQTEEDERLRAFTEAMTRYAPBGDPPRPBYDLELITSCSNVSVALGPG 2820
Db 2761 DLVVISSESQTEEDERLRAFTEAMTRYAPBGDPPRPBYDLELITSCSNVSVALGPG 2820
Qy 2821 RRRYVYTRDPTTPIAARAMEVTHSPVNSWLNIIQYAPTIWARMVLMTHFESIIMAAQOT 2880
Db 2821 RRRYVYTRDPTTPIAARAMEVTHSPVNSWLNIIQYAPTIWARMVLMTHFESIIMAAQOT 2880
Qy 2881 LDONLNFEMYGAVYSPLDPAIIBRLHGLDAFSLHTYTPHELTRVASALRKLGAAPLR 2940
Db 2881 LDONLNFEMYGAVYSPLDPAIIBRLHGLDAFSLHTYTPHELTRVASALRKLGAAPLR 2940
Qy 2941 AKMSRAAAYASLISGGRAAVCGRYLFNMAVTKLKTLPPLPARLLDSSWPTVAGAGG 3000
Db 2941 AKMSRAAAYASLISGGRAAVCGRYLFNMAVTKLKTLPPLPARLLDSSWPTVAGAGG 3000
Qy 3001 DIYHSVBARPRLLIFGLLLFVGGLFLLPAR 3033
Db 3001 DIYHSVBARPRLLIFGLLLFVGGLFLLPAR 3033
Qy 3033 DIYLCVLRARPRLLISLLISVGLFLLPAR 3033
Db 3033 DIYLCVLRARPRLLISLLISVGLFLLPAR 3033
```

```
RESULT 15
0909A9 PRELIMINARY; PRT: 3032 AA.
AC 0909A9;
DT 01-MAY-2000 (Tremblrel. 13, Created)
```

```
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DB 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DR Polypeptide.
OC Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OX NCBI_TaxID=11103;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=21361470; PubMed=11468731; DOI=10.1002/jmv.1073;
RA Kurhara C., Ishiyama N., Nishiyama Y., Fukushi S., Kageyama T.,
RA Kageyama K., Miura S.;
RT "Molecular characterization of hepatitis C virus genotype 2a from the
RT entire sequences of four isolates.";
RL J. Med. Virol. 64:466-475(2001).
DR EMBL: AF169003; AAF25611.1; -.
DR HSBP; O8YSL1; ICWX.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019078; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; Cytochrome_B5.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS5a.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_NS5a.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PsVlr.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; Viral_RdRP; 1.
DR SMART; SMO0487; DEXDc; 1.
DR PROSITE; PS00190; CYTOCHROME_C;
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polypeptide; Transmembrane.
SQ SEQUENCE 3032 AA; 329397 MW; BAF67F4730C124B9 CRC64;
```

```
Query Match 94.6%; Score 15249.5; DB 2; Length 3032;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 2847; Conservative 69; Mismatches 116; Indels 1; Gaps 1;
Qy 1 MSTNPKPQRTKRTNTRRPQDVKKPPGGQIVGVYLLPRGPRIGVRASTRKTSERSQPRG 60
Db 1 MSTNPKPQRTKRTNTRRPQDVKKPPGGQIVGVYLLPRGPRIGVRASTRKTSERSQPRG 60
Qy 61 RRQPIPKRRRSTGKSGWGPYWPPLYXGNEGIGMAGWLLSPRGSRRSPWPNRRRSRVNG 120
Db 61 RRQPIPKRRRSTGKSGWGPYWPPLYXGNEGIGMAGWLLSPRGSRRSPWPNRRRSRVNG 120
```

QY 121 KYIDLTGCGFADLMGXIPIVVGAPLGGVVARALAHGVRLVEDGVNPFATGNLPGCSFSIFILIA 180
DB 121 KYIDLTGCGFADLMGXIPIVVGAPLGGVVARALAHGVRLVEDGVNPFATGNLPGCSFSIFILIA 180
QY 181 ILSCITTPVSAAEVKNISGTVMYVNDCTNDSITWQAOAAVLHVPQGVPCPEKVGNAQOCMI 240
DB 181 ILSCITTPVSAAEVKNISGTVMYVNDCTNDSITWQAOAAVLHVPQGVPCPEKVGNAQOCMI 240
QY 241 PVSBNVAVORPQALVQGLRTHIDMVMYMATLCSALYVGLCGGVMIAOMFVSRQOHMF 300
DB 241 PVSBNVAVORPQALVQGLRTHIDMVMYMATLCSALYVGLCGGVMIAOMFVSRQOHMF 300
QY 301 VODNCSTIPIGTITGHRMAMDMMMNMSPTATMILYAMRVPEYIIDISGAMGVMPGLA 360
DB 301 VODNCSTIPIGTITGHRMAMDMMMNMSPTATMILYAMRVPEYIIDISGAMGVMPGLA 360
QY 361 YFSMOGAMAKVVVILLLAAGVDAARTHTVGSAAOTTGRLTSLPDMGPROKIQLVNTNGSM 420
DB 361 YFSMOGAMAKVVVILLLAAGVDAARTHTVGSAAOTTGRLTSLPDMGPROKIQLVNTNGSM 420
QY 421 HINRPAALNCNDSLHGFIALSLFYTHSFNSSGCPERMASACRSIEARRVMGALOYEDNVTN 480
DB 421 HINRPAALNCNDSLHGFIALSLFYTHSFNSSGCPERMASACRSIEARRVMGALOYEDNVTN 480
QY 481 PEDMKPYCMHYPKPCCKXIVPAGSVGCVCTPSPVNVGTTRDLGAPYTWGENEDVPL 540
DB 481 PEDMKPYCMHYPKPCCKXIVPAGSVGCVCTPSPVNVGTTRDLGAPYTWGENEDVPL 540
QY 541 LNSTRPPLGSMFGCTMNNSSGYTKTCGAPPCRTBADFNASTDLCTDCCFRKHPPTYLK 600
DB 541 LNSTRPPLGSMFGCTMNNSSGYTKTCGAPPCRTBADFNASTDLCTDCCFRKHPPTYLK 600
QY 601 CGSGGWLTPRCILIDYPIRLMHPCTVNVYTIPIKIMYVGVGEBRLTPACNFTGDRCNLED 660
DB 601 CGSGGWLTPRCILIDYPIRLMHPCTVNVYTIPIKIMYVGVGEBRLTPACNFTGDRCNLED 660
QY 661 RRSQSLPLHSTHSTEMALIPCSYSDLPALSTGLLHONIVDVOMPYGSPALTYCIYVM 720
DB 661 RRSQSLPLHSTHSTEMALIPCSYSDLPALSTGLLHONIVDVOMPYGSPALTYCIYVM 720
QY 721 EWWVILFLLLADARVACACLMMLILGOAEAALEKVLIIHAASAASNGEFLYFVIFVAAV 780
DB 721 EWWVILFLLLADARVACACLMMLILGOAEAALEKVLIIHAASAASNGEFLYFVIFVAAV 780
QY 781 YIKGRVPLATYSLTGLMSFSLILIALPOQAYAYDASVHGQIGALLVMTLFTLTPGYK 840
DB 781 YIKGRVPLATYSLTGLMSFSLILIALPOQAYAYDASVHGQIGALLVMTLFTLTPGYK 840
QY 841 TLLSRFLMWLCYLLTLGEMVQEBNAPPMQOVGRGRDGIWAVALFYPGVFODITKMLIANL 900
DB 841 TLLSRFLMWLCYLLTLGEMVQEBNAPPMQOVGRGRDGIWAVALFYPGVFODITKMLIANL 900
QY 901 GPAYLLKGLATRVYFVRHALLRMCCTMARHLAGRGYVOMALLALGRMTGTIYDHLTPM 960
DB 901 GPAYLLKGLATRVYFVRHALLRMCCTMARHLAGRGYVOMALLALGRMTGTIYDHLTPM 960
QY 961 SDMAASGLRDLAFAVEPIIFSPMEKKVIVMGAETAACGDIHLGLPVSARLAGEVILLGPAD 1020
DB 961 SDMAASGLRDLAFAVEPIIFSPMEKKVIVMGAETAACGDIHLGLPVSARLAGEVILLGPAD 1020
QY 1021 GYTSKGMSLIARITAYAOOTRGLIGTIYVSMGTGRDKTEOAGRIQVLTSTYTGOSFLGTISIG 1080
DB 1021 GYTSKGMSLIARITAYAOOTRGLIGTIYVSMGTGRDKTEOAGRIQVLTSTYTGOSFLGTISIG 1080
QY 1081 VLMYTVHAGAKNTLAGSRPVTOMYSASAAGDLVGWMPSPGTSLEBCTCGADVLIVLTBN 1140
DB 1081 VLMYTVHAGAKNTLAGSRPVTOMYSASAAGDLVGWMPSPGTSLEBCTCGADVLIVLTBN 1140
QY 1141 ADVIPARRRGDGRGALLSPRLSTLKSGSGGPLYCPRGHAGVFPRAAVCSRGVANSIDPI 1200
DB 1141 ADVIPARRRGDGRGALLSPRLSTLKSGSGGPLYCPRGHAGVFPRAAVCSRGVANSIDPI 1200
QY 1140 ADVIPARRRGDGRGALLSPRLSTLKSGSGGPLYCPRGHAGVFPRAAVCSRGVANSIDPI 1199

QY 1201 PVETLIDIVTRSPFSNDNSTPPAVPQTVGYVLYLHAPTGSGSKSTKVAVAAAGYKVLVLP 1260
DB 1201 PVETLIDIVTRSPFSNDNSTPPAVPQTVGYVLYLHAPTGSGSKSTKVAVAAAGYKVLVLP 1260
QY 1261 SVAATLFGAYILSKAHGINPNIIRTVGTGAPITTYSTYKFLADGGCAGAYIICD 1320
DB 1261 SVAATLFGAYILSKAHGINPNIIRTVGTGAPITTYSTYKFLADGGCAGAYIICD 1320
QY 1321 ECHAVDSYTIIGIGVULDOAEFAGVRLTVLTAATPGSVTPPHNIEBALQOEBEIPRY 1380
DB 1321 ECHAVDSYTIIGIGVULDOAEFAGVRLTVLTAATPGSVTPPHNIEBALQOEBEIPRY 1380
QY 1381 GRAIPLSYIKGGRHLIFCHSKKKCDBELAAALRGMLNVAAYRRGLDVSITPQGVVVVA 1440
DB 1381 GRAIPLSYIKGGRHLIFCHSKKKCDBELAAALRGMLNVAAYRRGLDVSITPQGVVVVA 1440
QY 1441 TDALMTGYTGPDSVIDCNVAATQVVDPSLDPTITTOIVPQDAVSRORSRGTGRGL 1500
DB 1441 TDALMTGYTGPDSVIDCNVAATQVVDPSLDPTITTOIVPQDAVSRORSRGTGRGL 1500
QY 1501 GIYRVVSTGERASGMFDSVULCECYDACAAMVELTPSETVRLRAYFNTPGLPVQCQDHL 1560
DB 1501 GIYRVVSTGERASGMFDSVULCECYDACAAMVELTPSETVRLRAYFNTPGLPVQCQDHL 1560
QY 1561 FWEAVFTGLTHIDAFHLSQTKOSGENFAYLTAQATVCARAKAPPSMDVMKCLTRLKP 1620
DB 1561 FWEAVFTGLTHIDAFHLSQTKOSGENFAYLTAQATVCARAKAPPSMDVMKCLTRLKP 1620
QY 1621 TLVGPPLLYRLGVTNEVTLTHPVTKYIATCMQADLEVMSTWVLAGVILAAVAYCIA 1680
DB 1621 TLVGPPLLYRLGVTNEVTLTHPVTKYIATCMQADLEVMSTWVLAGVILAAVAYCIA 1680
QY 1681 TGCVCILIGRLHNOAAVAPDKEVLYEAPDEMEBCASPAALIEBQRIAEMLRSKIOGL 1740
DB 1681 TGCVCILIGRLHNOAAVAPDKEVLYEAPDEMEBCASPAALIEBQRIAEMLRSKIOGL 1740
QY 1741 QOASQAOADIOPVQASWPKEQOFKAKMMNFIISQIYLAGLSTPGNPAVASMMAFSA 1800
DB 1741 QOASQAOADIOPVQASWPKEQOFKAKMMNFIISQIYLAGLSTPGNPAVASMMAFSA 1800
QY 1801 LTPSLSTSTILLNLLGMLASQIAPPAAGTGFVVSGLVGAAGSISGLKVLVDILLAGY 1860
DB 1801 LTPSLSTSTILLNLLGMLASQIAPPAAGTGFVVSGLVGAAGSISGLKVLVDILLAGY 1860
QY 1861 AGISGALVAFKIMSEKSMENVNLLPGIISPGALVYGVICAAILRRHVGBGAOVOM 1920
DB 1861 AGISGALVAFKIMSEKSMENVNLLPGIISPGALVYGVICAAILRRHVGBGAOVOM 1920
QY 1921 NRLIPASRGHNVAPTHVVTESDASQRYTOLLGSLTISLRLRHNMTEDCPIPCGSM 1980
DB 1921 NRLIPASRGHNVAPTHVVTESDASQRYTOLLGSLTISLRLRHNMTEDCPIPCGSM 1980
QY 1981 LRDVMDWYCTILLTDEKMLJTSKLPKMPGLPFVSCQKYGKVMAGTGIMTTRPCPGANIS 2040
DB 1981 LRDVMDWYCTILLTDEKMLJTSKLPKMPGLPFVSCQKYGKVMAGTGIMTTRPCPGANIS 2040
QY 2041 GNVRLGSRIRIGPKTCANITWQTFPINCYTBQOCYKPAKPNKVALIMVAASEYAEVTOH 2100
DB 2041 GNVRLGSRIRIGPKTCANITWQTFPINCYTBQOCYKPAKPNKVALIMVAASEYAEVTOH 2100
QY 2101 GSYYHITGLTJDNMLVPCQLSPSEFFSWVDGQVHRFPKPPKPFREDEVSFCVGLNSPV 2160
DB 2101 GSYYHITGLTJDNMLVPCQLSPSEFFSWVDGQVHRFPKPPKPFREDEVSFCVGLNSPV 2160
QY 2161 GSQPCDEPDTVLMSMLTDPSSHITAETAARLARLARGSPSEBASSASQSLAPSLATCT 2220
DB 2161 GSQPCDEPDTVLMSMLTDPSSHITAETAARLARLARGSPSEBASSASQSLAPSLATCT 2220
QY 2221 THGKAYVDVMDANLFMGQVYTRIESGSKVYVJDSLDPMVEERSLPESIPSEYMLPKR 2280
DB 2221 THGKAYVDVMDANLFMGQVYTRIESGSKVYVJDSLDPMVEERSLPESIPSEYMLPKR 2280
QY 2281 FPPALPAWARPDPYNPLVESWKRDPYQPAIVAGCALPPEPRKTPPTPPRRRTVGLSEDSI 2340

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2005, 15:50:12 ; Search time 65 Seconds
(without alignments)
3483.240 Million cell updates/sec

Title: US-09-980-559-2

Perfect score: 16120

Sequence: 1 MSTNPKPQKRTKRNTRRPQ.....LTFGLLLFVGVLFLPAR 3033

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	15812	98.1	3033	1	US-07-925-695-5	Sequence 5, Appl
2	13966	86.6	3033	1	US-07-925-695-8	Sequence 8, Appl
3	13881	86.1	3033	1	US-07-925-695-9	Sequence 9, Appl
4	12027.5	74.6	3010	4	US-09-539-601-3	Sequence 3, Appl
5	12010.5	74.5	3010	4	US-09-539-601-21	Sequence 21, Appl
6	12001.5	74.5	3010	4	US-09-539-601-27	Sequence 27, Appl
7	11984.5	74.3	3010	1	US-08-324-977-2	Sequence 2, Appl
8	11984.5	74.3	3010	1	US-08-324-977-14	Sequence 14, Appl
9	11984.5	74.3	3010	2	US-08-384-616-2	Sequence 2, Appl
10	11984.5	74.3	3010	2	US-08-384-616-14	Sequence 14, Appl
11	11984.5	74.3	3010	2	US-08-904-686A-2	Sequence 2, Appl
12	11984.5	74.3	3010	2	US-08-904-686A-14	Sequence 14, Appl
13	11984.5	74.3	3010	2	US-09-315-850-2	Sequence 2, Appl
14	11984.5	74.3	3010	4	US-09-315-850-14	Sequence 14, Appl
15	11984.5	74.3	3010	4	US-09-539-601-33	Sequence 33, Appl
16	11973	74.3	3012	3	US-08-811-566-2	Sequence 2, Appl
17	11973	74.3	3012	3	US-09-034-756-2	Sequence 2, Appl
18	11966	74.2	3011	3	US-09-014-416-1	Sequence 1, Appl
19	11966	74.2	3011	4	US-09-952-572-9	Sequence 9, Appl
20	11958.5	74.2	3010	4	US-09-014-416-3	Sequence 3, Appl
21	11947	74.1	3011	1	US-08-453-552-2	Sequence 2, Appl
22	11947	74.1	3011	2	US-08-710-637-2	Sequence 2, Appl
23	11945	74.1	3011	5	PCT-US93-00907-2	Sequence 2, Appl
24	11945	74.1	3011	3	US-09-014-416-5	Sequence 5, Appl
25	11933	74.0	3011	3	US-08-850-328-1	Sequence 1, Appl
26	11922	74.0	3011	1	US-08-440-103-36	Sequence 36, Appl
27	11922	74.0	3011	1	US-08-440-542-36	Sequence 36, Appl

28	11922	74.0	3011	1	US-07-910-760-10	Sequence 10, Appl
29	11922	74.0	3011	1	US-08-440-519-10	Sequence 10, Appl
30	11922	74.0	3011	1	US-08-231-368-36	Sequence 36, Appl
31	11922	74.0	3011	1	US-08-440-210-36	Sequence 36, Appl
32	11922	74.0	3011	3	US-09-388-874-2	Sequence 2, Appl
33	11922	74.0	3011	3	US-09-046-604-36	Sequence 36, Appl
34	11922	74.0	3011	4	US-08-440-549-10	Sequence 10, Appl
35	11922	74.0	3011	4	US-09-916-359-2	Sequence 2, Appl
36	11915	73.9	3011	1	US-08-188-281B-1	Sequence 1, Appl
37	11915	73.9	3011	1	US-08-453-552-1	Sequence 1, Appl
38	11915	73.9	3011	2	US-08-710-637-1	Sequence 1, Appl
39	11915	73.9	3011	5	PCT-US93-00907-1	Sequence 1, Appl
40	11915	73.9	3011	5	PCT-US94-07280-1	Sequence 1, Appl
41	11915	73.9	3011	5	PCT-US95-01087-1	Sequence 1, Appl
42	11900	73.8	3011	3	US-08-811-566-20	Sequence 20, Appl
43	11900	73.8	3011	3	US-09-034-756-20	Sequence 20, Appl
44	11857	73.6	2995	3	US-08-444-818-138	Sequence 138, App
45	11828	73.4	3011	5	PCT-US91-02225-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-07-925-695-5
; Sequence 5, Application US/07925695
; Patent No. 5428145
; GENERAL INFORMATION:
; APPLICANT: OKAMOTO, Hiroaki
; APPLICANT: NAKAMURA, Tetsuo
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOME,
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
; TITLE OF INVENTION: DETECTION SYSTEMS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beveridge, DeGrandi, Weillacher & Young
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: US
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/925,695
; FILING DATE: 19920807
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 287402/91
; FILING DATE: 09-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 360441/91
; FILING DATE: 05-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Weillacher, Robert G.
; REGISTRATION NUMBER: 20,531
; REFERENCE/DOCKET NUMBER: 06/87-48009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2811
; TELEFAX: (202) 659-1462
; TELEX: WUI 64470
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3033 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; US-07-925-695-5
Query Match 98.1%; Score 15812; DB 1; length 3033;

Best Local Similarity 97.7%; Pred. No. 0;
Matches 2964; Conservative 31; Mismatches 38; Indels 0; Gaps 0;

[illegible]

Db	1021	GYSKMSLAPITAYAQORGLGTTIVSMSTGDKTEQAGEIEVLSVTWQSFLGTTISG	1080
Qy	1081	VLMTVYHAGNKTLAGSRGPVTQWYSSAEGDLVGMPSPEGTKSLEPCTCGAVDLYLVTRN	1140
Db	1081	VLMTVYHAGNKTLAGSRGPVTQWYSSAEGDLVGMPSPEGTKSLEPCTCGAVDLYLVTRN	1140
Qy	1141	ADVPARRRGKRALISPRPLSTLKSSSGSPVLCPRGHAVGVRAVCSGVAKSIDFI	1200
Db	1141	ADVPARRRGKRALISPRPLSTLKSSSGSPVLCPRGHAVGVRAVCSGVAKSIDFI	1200
Qy	1201	PVELTLDIVTSPTRESDSNTPPAVPOQTQOVGLHAFTPSGSKTKVPVAAQGVKVLVLP	1260
Db	1201	PVELTLDIVTSPTRESDSNTPPAVPOQTQOVGLHAFTPSGSKTKVPVAAQGVKVLVLP	1260
Qy	1261	SVAATLFGAYLSTAHGINPNIRTVRTVTGADITYSTYCKFLADGCGAGAYDIIICD	1320
Db	1261	SVAATLFGAYLSTAHGINPNIRTVRTVTGADITYSTYCKFLADGCGAGAYDIIICD	1320
Qy	1321	ECHAVDSFTLLIGITVDLQAEFAVRLTVLATAPEPSVTTPHNIIEFVALGOGGEIPEY	1380
Db	1321	ECHAVDSFTLLIGITVDLQAEFAVRLTVLATAPEPSVTTPHNIIEFVALGOGGEIPEY	1380
Qy	1381	GRAIPLSTYIKGRHLIFCHSKKCCDELAAALRGGLSVAAYYRGJADVSVPJTGQDVVVA	1440
Db	1381	GRAIPLSTYIKGRHLIFCHSKKCCDELAAALRGGLSVAAYYRGJADVSVPJTGQDVVVA	1440
Qy	1441	TDALMTGYTGDPSVIDCNVAVTQVVDPSLDPTFTITTOJVPDQAVSRQGRGTGRGL	1500
Db	1441	TDALMTGYTGDPSVIDCNVAVTQVVDPSLDPTFTITTOJVPDQAVSRQGRGTGRGL	1500
Qy	1501	GIYRYVSTGERASGMPDSVULCECYDAGAANYELTPESETTYRLAAYFNTPELPCODHLE	1560
Db	1501	GIYRYVSTGERASGMPDSVULCECYDAGAANYELTPESETTYRLAAYFNTPELPCODHLE	1560
Qy	1561	FMEAVFTGLTHIDAHPLISQTKOSENFAYLTAQATYCARAKAPPSMDWMMKCLTRPKP	1620
Db	1561	FMEAVFTGLTHIDAHPLISQTKOSENFAYLTAQATYCARAKAPPSMDWMMKCLTRPKP	1620
Qy	1621	TLVGPETPLLYLGSVTNEBVTLTHPEVTXYKIAICMQADLEVMTSTWVLAGVLAALVAAYCLA	1680
Db	1621	TLVGPETPLLYLGSVTNEBVTLTHPEVTXYKIAICMQADLEVMTSTWVLAGVLAALVAAYCLA	1680
Qy	1681	TGCYCIIIGRLHINQAVVAPPKEXVLYEAFDEMECASBAALIEGQRIAEMLKSKIQGLL	1740
Db	1681	TGCYCIIIGRLHINQAVVAPPKEXVLYEAFDEMECASBAALIEGQRIAEMLKSKIQGLL	1740
Qy	1741	QOASKOAOADIPYQOASMPKVEQGMKMMNFISIGIOYLACTSLTPGNPAVASMMAFSA	1800
Db	1741	QOASKOAOADIPYQOASMPKVEQGMKMMNFISIGIOYLACTSLTPGNPAVASMMAFSA	1800
Qy	1801	LTPSLPSTSTLLNLILGMLASQIAPAPGATGFVSGLVGAAGVSGIGLKVLDIILGYG	1860
Db	1801	LTPSLPSTSTLLNLILGMLASQIAPAPGATGFVSGLVGAAGVSGIGLKVLDIILGYG	1860
Qy	1861	AGISGALVAFKINGEKSPMEDVNNLPGIISPAALVGVICAAILRRHVGEGEAVQM	1920
Db	1861	AGISGALVAFKINGEKSPMEDVNNLPGIISPAALVGVICAAILRRHVGEGEAVQM	1920
Qy	1921	NRLIAFASRGHVAPTHVYTESDASQRTQLGSLTTISLLRRLHMNTTEBCPLPCSSSW	1980
Db	1921	NRLIAFASRGHVAPTHVYTESDASQRTQLGSLTTISLLRRLHMNTTEBCPLPCSSSW	1980
Qy	1981	LRDVWDWCTLLTFPKMWLTSKLPPKMGPLPFVSCQGYKGVMAGTGIMTTRCCGANIS	2040
Db	1981	LRDVWDWCTLLTFPKMWLTSKLPPKMGPLPFVSCQGYKGVMAGTGIMTTRCCGANIS	2040
Qy	2041	GNVRLGSMRITGPKTCNNIWQGTPEPINCYTEGQCVCPKAPNFKXIAIMRWASBYAEVTOH	2100
Db	2041	GNVRLGSMRITGPKTCNNIWQGTPEPINCYTEGQCVCPKAPNFKXIAIMRWASBYAEVTOH	2100
Qy	2101	GSYHYITGLTNDNLKVPQCLPSPEFESWDOVQIHRAPPTYKPPRDEVESTCYGLNSVY	2160
Db	2101	GSYHYITGLTNDNLKVPQCLPSPEFESWDOVQIHRAPPTYKPPRDEVESTCYGLNSVY	2160

Qy	481	PEDMRPVCMHYPRROCGVVS	SAKTVCGPVYCFPSPVWVGTDR	LGAPYTTMGENETVFL	540			
Db	481	DGDMRPVCMHYPRPCGIV	PARVIVCGPVYCFPSPVWVGTDR	KGVPTTVMGENETVFL	540			
Qy	541	LNSTRPLSGMFCCTMWNSS	GVYKTCGAPCPCRTAD	FMASTDLCPTCCFKHEDTYLK	600			
Db	541	LNSTRPRRGAMPCCTMMNG	GTGFKTCGAPPCIRKDVN	STIDLCPDCCFKHDAITYLK	600			
Qy	601	CGSGPWLT	PPCLIDYPRLMHPCTVYNTI	FKIRMTYQGVBEHRLTAA	CNFTGRGRCNLED	660		
Db	601	CGAGPWLT	PPCLVDYPRLMHPCTVYNTI	FFARMYGVBEHRSFMA	CNFTGRGRCNLED	660		
Qy	661	RDRSOLSP	LLHSTTEMAALPCSSVDL	PALSTGLH	HOINVDVOMYGLSALPKYIVRN	720		
Db	661	RDRQOQSP	LLHSTTEMAAVLPCSSVDL	PAJSTGLH	HOINVDVQVLYGLSALPKYIVRN	720		
Qy	721	EWVILFL	LLADARVACIMLMLIL	GOAEBALEKULVIL	MAASACNGFLFVIFVYAAW	780		
Db	721	EWVILFL	LLADARICACIMMLIIT	GOBALEKULIIL	HSASASANGPLMFFIFVYAAW	780		
Qy	781	YIKGRVVP	PLATYSLTGLMFS	SLILLLALPOOAYAY	ADASVHGQIGAAALVMITL	PLUTPGYK	840	
Db	781	YLKGRVVP	AVATYSVLGLMFS	FLIVYALPOOAYAL	MAEGBEIGLAIIVYIISIFL	PLTPAYK	840	
Qy	841	TLISRP	LMWLCYLLTTLG	EAUVOMAPMPOVGR	GGRIIMANAIFPVGPVPIT	TKMLAVL	900	
Db	841	TLISRS	TWMLSTYLVLAEAO	IQOWPPELREGRBOIIM	AVIIMHPLVEFTVM	LALIL	900	
Qy	901	GPAVLL	KGALTTRVPYFV	RAHALLRMCTMAR	HLAGRYVQOMLAL	LGWGTGYIYDHLTPM	960	
Db	901	GPAVLL	KASLLRIPYF	RAHALLRVCTL	VKHLAGARIYOMLLIT	ITGRWGTGYIYDHLSP	960	
Qy	961	SDMAASL	RLDAAVBERPIFSP	MEKKYIYWGAETACGD	ILHGLPVSKRLGRV	VLBPAD	1020	
Db	961	STMAAQ	RLDLAAVEBVPF	MEKKYIYWGAEYACG	ILHGLPVSARLGR	EVLLBPAD	1020	
Qy	1021	GYTSKGM	SLAPITAAVACQ	TRGLLGTIVY	SMGRDTEBAGRIQVLS	TVTSPFGTSSISG	1080	
Db	1021	GYTSKGM	KLAPITATYI	QOTRGLLGAIVSL	GRDKRDEQAGOVQLSV	TVTFLGTSISG	1080	
Qy	1081	VLMTVY	HAGNKTLAGSR	BPVTOMYSABEGL	VGMPSPPGTSLEP	CTCGAVDLYLVRN	1140	
Db	1081	VLMTVY	HAGNKTLAGK	BPVTOMYSABGL	VGMPSPPGTS	LDPCGAVDLYLVRN	1140	
Qy	1141	ADVPRAR	RDDKXGALLSP	RLSTLKGSSGG	PVLCPGRHACV	VEFAAACSNGVAKSIDFI	1200	
Db	1141	ADVPRAR	KDDRKAGALLSP	RLSTLKGSSGG	PVLCSRGHAGV	LFPAACACAGVAKSIDFI	1200	
Qy	1201	PVELT	LDIVTRSPIF	FSDNSTRPAPAPQ	TYQYGYLHAPTS	SGSKSTKVYAAAGYKVLVLNP	1260	
Db	1201	PVESL	DVAATTPSP	SDNSTRPAPAPQ	TYQYGYLHAPTS	SGSKSTKVPAABAOSGYKVLVLNP	1260	
Qy	1261	SVAAT	LCFGAYLSK	AGINPNIRTVGR	IVYTTGAPIT	ITYSYGKEFLADGGCAGAYDIICD	1320	
Db	1261	SVAAT	LCFGAYMSK	AGINPNIRTVGR	IVYTTGDSI	ITYSYGKEFIADGGCAAGAYDIICD	1320	
Qy	1321	EGHAV	VDSTTLIGITV	LDQAE	TAGVRLVLAATAP	PGSGTTPHPHIEBVAL	GOGEGEIPFY	1380
Db	1321	EGHSV	VDATTLIGITV	LDQAE	TAGVRLVLAATAP	PGVITTPHSHIEBVAL	GHHGEIIPFY	1380
Qy	1381	GRAIP	PLSYIRGGRHLIF	CHSKKKKCD	ELAAALNGMGLNS	AVAYYRGIDVAVIPTOGD	VVVVA	1440
Db	1381	GKAIP	PLAFIRGGRHLIF	CHSKKKKCD	ELAAALNGMGNA	ANAYYRGIDVAVIPTOGD	VVVVA	1440
Qy	1441	TDALMT	GYTGD	PDSVIDCN	AVATVOVDFSL	DPITTTITQIV	PODAYSRSORRGRTGRL	1500
Db	1441	TDALMT	GYTGD	PDSVIDCN	AVATVOVDFSL	DPITTTITQIV	PODAYSRSORRGRTGRL	1500
Qy	1501	GIYVY	VTGGRASG	MPSVVLGC	CYDAGAAWELP	SETTVYLRAYE	YNTPGLPYCQDHL	1560
Db	1501	GVAYV	SSGGRPSG	MPSVVLGC	CYDAGAAWELP	SETTVYLRAYE	YNTPGLPYCQDHL	1560

QY	1561	FWEAVFGLTHIDHNFISQTKOSGSENAVYLTAQATCAACAKAPBPBDDMMKCTFELKP	1620
Db	1561	FWEAVFGLTHIDHNFISQTKOSGSENAVYLTAQATCAACAKAPBPBDDMMKCTFELKP	1620
QY	1621	TLVAPTPLYLXGSVTMEVLLTHFVTKYIATCMQADIEVMSTVWVLGCVLAAVAACLA	1680
Db	1621	TLVAPTPLYLXGAVTMEVLLTHFVTKYIATCMQADIEVMSTVWVLGCVLAAVAACLA	1680
QY	1661	TGCYCITIGRLHINORAVAPDKVELYAPFPMECASRALITBEGONIAEMTKSIGLL	1740
Db	1661	TGCISITIGRLHINRVAVAPDKVELYAPFPMECASRAALITEGOMAEMLSKTIGLL	1740
QY	1741	QOASQOAOODIOPYQOASMPKVECFQWAGHMNFISIGIYTLAGLSTLPGNPAVASMAFSA	1800
Db	1741	QOARROAQODIOPALQOSMPKLEQWAGHMNFISIGIYTLAGLSTLPGNPAVASMAFSA	1800
QY	1801	LTPSLPSTITLLNLIGWLASQIAPPAAGATGFVYSGVLGAAVSGITLGVLYDLIAGYG	1860
Db	1801	LTPSLPSTITLLNLIMGWLASQIAPPAAGATGFVSGVLGAAVSGITLGVLYDLIAGYG	1860
QY	1861	AGISGALVAFKIMGSEKXPMEDEVNLLPGJISPCALVGVYCAILRRHVGEAGVQM	1920
Db	1861	AGISGALVAFKIMGSEKPYEDVNNLPALISPALVGVYCAILRRHVGEAGVQM	1920
QY	1921	NRLJAFASRGNHVAPTHYVYESDASQRYTOLLSITTSLLRLHNNITEDCPIPCGGSW	1980
Db	1921	NRLJAFASRGNHVAPTHYVYESDASQRYTOLLSITTSLLRLHNNITEDCPIPCGGSW	1980
QY	1981	LRDYWDWVCITLTPDFKNNLTSKLEPKPKGJLPEVSCQGYGVWAGTIGMTTRCCGANIS	2040
Db	1981	LQDIDWDCSLTPDFKNNLSKLEPKPKGJLPEVSCQGYGVWAGTIGMTTRCCGANIS	2040
QY	2041	GNVVLGSMRITGPXTCKNNIMQGETPINCYTEGQCVKPAFNFKYAIWRVASEAYEATOH	2100
Db	2041	GHVAMGMVTKITGPXTCKNLIMQGETPINCYTEGQCVKPPRPYKXAIWRVASEAYEATOH	2100
QY	2101	GSYHITIGLTLDNLKVCQOLBPSEFPFSDVQVQHRAPRTPKPFPRBDESGCVLNSFV	2160
Db	2101	GSFSYVGLTSDNLKVCQVPAPBFPSWQDVQIHRAPVBPFRBDEVTGVLNSFV	2160
QY	2161	GSOLPCDEPPTDVLMSTLTPSHITAEATARLARLARSPESEASSASQLASPILRATCT	2220
Db	2161	GSOLPCDEPPTVELAMTLDPHSITAEATARLARLARSPEQASSASQLASPILKATCT	2220
QY	2221	THGKAYDVADVANLPMGQDVTRIESGSKVVVLDLSDPMVEERSDLBSPISEYMLPKR	2280
Db	2221	THKTAAYDCDWDANLPMGQDVTRIESSKVIVLDSLSPMEVEBDRBSPSEYMLIKRK	2280
QY	2281	FPPALPMARADYNPVLVESWKRDRDYOPATVAGALPPRKTPTPPRRRRVTSSESI	2340
Db	2281	FPPALPMARADYNPVLITWKRCYEPYVLCALPPTPTTPPPRRRRRAKVLTDNV	2340
QY	2341	GDALQOLAKISFGQPPPSGDSGLSTGAGADSQTPPEDELATSETGISISSMPLSEBLG	2400
Db	2341	EGVLRREMADKYLSPLQDNNDSGHSTGADTGADIYQOQSDDETAABEASLSMPLSEBPG	2400
QY	2401	DPLELEPQVEQPPRQGVAPRGSDSQSWTSCSESDSVVCCSMYSMTGALLTPRCPBE	2460
Db	2401	DPDLEFEFVSGAPSEGEVIDSDSWSVTSQEBVSIVCCSMYSMTGALLTPRCPBE	2460
QY	2461	EKLPIPLNSNLTLYHNKVCYCTTKSASLRAKKCTPRMQVLDSYUDSVLMDIKLAASKV	2520
Db	2461	EKLPIPLNSNLTMYRPHNKVYSTTSRSASLRKKCTPRVQVLDLDAHYDSVLQDVRASAKV	2520
QY	2521	TARLTMEACQLTIPRHSARSKYFGAKEVFSLSGRAVNNIKSVWKDLLESEFPITTI	2580
Db	2521	SARLTVEACALTPRHSAKRYFGAKEVFSLSRRAVNNIRSWENLLEOHPIDITTI	2580
QY	2581	MAKNEVFCVDPDTKGKKAAALIVYPDIGVRCCEMALYDIQOKLPOAMGASYGFQVSPA	2640
Db	2581	MAKNEVFCIDPDTKGKKARLIVYPDIGVRCCEMALYDIQOKLPKAIMGSIYGFQVSPA	2640
QY	2641	QVVELLKAMAEKQDPMGFSDYTRCFQSTVERDIRTEESIVYRACSJPEEAHTAHSLTE	2700

```

Db      2641 ERVDFLKAMGSKDPMGFSDYTRCFDSTVTERDITBESIYQACSLPQEARVHSLRE 2700
Qy      2701 RLYVGGPMNSKQGTGGRCAASGVLTTSMGNTTTCYKALAAAGIATMVLVCGD 2760
Db      2701 RLYVGGPMNSKQGTGGRCAASGVLTTSMGNTTTCYKALAAAGIATMVLVCGD 2760
Qy      2761 DLVVISQGTEDENLAFTEAMTRYSAPEGDPREPEXDELITSCSSNVVALGPOG 2820
Db      2761 DLVVISQGTEDENLAFTEAMTRYSAPEGDPREPEXDELITSCSSNVVALGPOG 2820
Qy      2821 RRRYVLTDPPTPIABAEVTVHSPVNSMIGNIIOYAPTIVARVLTMTFFSILMAODT 2880
Db      2821 RRRYVLTDPPTPIABAEVTVHSPVNSMIGNIIOYAPTIVARVLTMTFFSILMAODT 2880
Qy      2881 LDONLNFEMYGAVYSVPLDPAIIRLHGLDAFSLHTTTPHELTIVASAKRLGAPPLR 2940
Db      2881 LDONLNFEMYGAVYSVPLDPAIIRLHGLDAFSLHTTTPHELTIVASAKRLGAPPLR 2940
Qy      2941 AMKSRARAVASLISRGRAAVGRLFMNAVTKLKLPLPARLDDISSWFTVAGGG 3000
Db      2941 AMKSRARAVASLISRGRAAVGRLFMNAVTKLKLPLPARLDDISSWFTVAGGG 3000
Qy      3001 DIHVSRRARPRLLFGLLLFVGVGLFLPLPAR 3033
Db      3001 DIHVSRRARPRLLFGLLLFVGVGLFLPLPAR 3033
Qy      3001 DIHVSRRARPRLLFGLLLFVGVGLFLPLPAR 3033
Db      3001 DIHVSRRARPRLLFGLLLFVGVGLFLPLPAR 3033

```

```

RESULT 3
US-07-925-695-9
; Sequence 9, Application US/07925695
; Patent No. 5428145
; GENERAL INFORMATION:
; APPLICANT: OKAMOTO, Hiroaki
; APPLICANT: NAKAMURA, Tetsuo
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOME,
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
; NUMBER OF INVENTIONS: 9
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beveridge, Degrandi, Weillacher & Young
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: US
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Releasee #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/925,695
; FILING DATE: 19920807
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 287402/91
; FILING DATE: 09-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 360441/91
; FILING DATE: 05-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Weillacher, Robert G.
; REGISTRATION NUMBER: 20,531
; REFERENCE/DOCKET NUMBER: 06/87-48009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2811
; TELEFAX: (202) 659-1462
; TELEX: WUI 64470
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3033 amino acids
; TYPE: AMINO ACID

```

```

; STRANDEDNESS: unknown
; TOPOLOGY: linear
US-07-925-695-9
Query Match      86.1%; Score 13881; DB 1; Length 3033;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 2534; Conservative 245; Mismatches 254; Indels 0; Gaps 0;

Qy      1 NSTNPKQKTRNTNRNRPODYKFPQGGQITVGGVYLLPFRGPRGLVGRATKTSERQOPG 60
Db      1 NSTNPKQKTRNTNRNRPODYKFPQGGQITVGGVYLLPFRGPRGLVGRATKTSERQOPG 60
Qy      61 RROP1PKDRSTGKSGKGPWPPLYGNGLGAGLLSPGSGPSPGPNDRSRNNG 120
Db      61 RROP1PKDRSTGKSGKGPWPPLYGNGLGAGLLSPGSGPSPGPNDRSRNNG 120
Qy      121 KYIDTLTCGFADLMGYIPVGAAPLGAVARALAHGVRLEEDGVNFAATGNLPGCSFIFLLA 180
Db      121 KYIDTLTCGFADLMGYIPVGAAPLGAVARALAHGVRLEEDGVNFAATGNLPGCSFIFLLA 180
Qy      181 LLSCTTPVSAABEVKNISGVMVTNDCTNDSTWQOAAVLHVPQVPCREKGNASQCVI 240
Db      181 LLSCTTPVSAABEVKNISGVMVTNDCTNDSTWQOAAVLHVPQVPCREKGNASQCVI 240
Qy      241 PVSFVAVQRPALTCGLRTHIDMVMSATLCSALYVGDLCGVMIAQMTIVSPQHMF 300
Db      241 PVSFVAVQRPALTCGLRTHIDMVMSATLCSALYVGDLCGVMIAQMTIVSPQHMF 300
Qy      301 VDDNCISYPGTTGHRMAMDMMNWSPTATMILAVAMRVEVIIDIISGAHWVFGLA 360
Db      301 VDDNCISYPGTTGHRMAMDMMNWSPTATMILAVAMRVEVIIDIISGAHWVFGLA 360
Qy      361 YFSMOGAMKVVVILLAAVUDARTTVGSGAAQTGRLTSLPDMQPRKQIOLVNNGS 420
Db      361 YFSMOGAMKVVVILLAAVUDARTTVGSGAAQTGRLTSLPDMQPRKQIOLVNNGS 420
Qy      421 HINRTALNCNDSLHTGFIASLYFTHSFNSSGPERMSACRSLEAFVGCALQYEDNVN 480
Db      421 HINRTALNCNDSLHTGFIASLYFTHSFNSSGPERMSACRSLEAFVGCALQYEDNVN 480
Qy      481 PEDMRPYCMHYPPROGCVSAKTVCGPVYCTPSPVAVGTTRDLGAPTYWGBNETDVL 540
Db      481 PEDMRPYCMHYPPROGCVSAKTVCGPVYCTPSPVAVGTTRDLGAPTYWGBNETDVL 540
Qy      541 INSTREPLGSGWFCCTMMSSGYTKTCGAPCCTRADFNASTLLCCTDPRKHPDPTYLK 600
Db      541 INSTREPLGSGWFCCTMMSSGYTKTCGAPCCTRADFNASTLLCCTDPRKHPDPTYLK 600
Qy      601 CGSGPWLTPRCIDIDPYRLMHYPCIVNYTIFKIRMYGVGVEHRLTAACNFTGRGRCNLD 660
Db      601 CGSGPWLTPRCIDIDPYRLMHYPCIVNYTIFKIRMYGVGVEHRLTAACNFTGRGRCNLD 660
Qy      661 RDRSOLSPLLHSTTEWAILPCSYSDLPALSTGLHLHONIIVDQFMYGSLPALTKYIVM 720
Db      661 RDRSOLSPLLHSTTEWAILPCSYSDLPALSTGLHLHONIIVDQFMYGSLPALTKYIVM 720
Qy      721 EWIYLLFLILADARVACACIMLILGOAFAALEKVIILHAASAGNSGLYVYIPVAVM 780
Db      721 EWIYLLFLILADARVACACIMLILGOAFAALEKVIILHAASAGNSGLYVYIPVAVM 780
Qy      781 YIKGRVPLATYSLGLMFSLLALPOQAAVADVASVGOIGALLVNTITFTLTPGYK 840
Db      781 YIKGRVPLATYSLGLMFSLLALPOQAAVADVASVGOIGALLVNTITFTLTPGYK 840
Qy      841 TLLSRFLMWLCYLLTLGEMVQEWAPPMQVGRGDDIIVAAVAFYPGVFDITKWLAVL 900
Db      841 TLLSRFLMWLCYLLTLGEMVQEWAPPMQVGRGDDIIVAAVAFYPGVFDITKWLAVL 900
Qy      901 GPAYILRLASLRIPIFVRHALRLVCTLVKHLAAGARYIOMLITIGRMGTGIYIDLSFL 960
Db      901 GPAYILRLASLRIPIFVRHALRLVCTLVKHLAAGARYIOMLITIGRMGTGIYIDLSFL 960
Qy      961 SDMAASGLRLDAVAVEPIIFSPEKVKVIWGAETAAAGDILHGLPSARLAGREVLLGPAD 1020

```

```

Db      961 STWAAQCLRDALALAVEPVFSPMEKVIWGAETVACGDIHGI.PVSARLGREVLGPAD 1020
Qy      1021 GYTSKGSLLAPITAYAQOORGLIGTVSMGTGDKTEQAGEIOVSTVOSFELGTSISG 1080
Db      1021 GYTSKGNLAPITAYAQOORGLIGTVSMGTGDKTEQAGEIOVSTVOSFELGTSISG 1080
Qy      1081 VLMVTYHAGNKTLAGSGPVYOMYSABEDLVGMPSPGTSLEPCTCGAVDLVLTNRN 1140
Db      1081 VLMVTYHAGNKTLAGSGPVYOMYSABEDLVGMPSPGTSLEPCTCGAVDLVLTNRN 1140
Qy      1141 ADVIPARRRDKGALLSPRPLSTLKSSSGGCVLCPRGHAGVGFRAAVCSRGVAKSIDFI 1200
Db      1141 ADVIPARRRDKGALLSPRPLSTLKSSSGGCVLCPRGHAGVGFRAAVCSRGVAKSIDFI 1200
Qy      1201 PVETLDIVTRSPRTSDNSTPRPAVPOQYOVGLHAHPTSGSGSTKVPVVAAYQYKVLVLP 1260
Db      1201 PVESLDIATRTPSSDNSAPPAVPOSYOVGLHAHPTSGSGSTKVPVVAAYQYKVLVLP 1260
Qy      1261 SVAATLGFAGYLSKXAGINPNIARTGVRTVTGAPITVSTYGFKLADGCGAGAYDIICD 1320
Db      1261 SVAATLGFAGYMSKXAGINPNIARTGVRTVTGDSITVSTYGFKLADGCGAGAYDIICD 1320
Qy      1321 ECHAVDSTTLIGITVDQAEFTAGVRLTVLATPESVTTTTPHNIEVALGOGGEIPIFY 1380
Db      1321 ECHSVADATTLIGITVDQAEFTAGVRLTVLATPESVTTTTPHNIEVALGHEGEIPIFY 1380
Qy      1381 GRAIPLEYIKGGRHLIFCHSKKKCDLAAALRGKGLSVAYRGLDVSV.PTQGDVVVVA 1440
Db      1381 GKALPDLPIFKGRHLIFCHSKKKCDLAAALRGKGLSVAYRGLDVSV.PTQGDVVVVA 1440
Qy      1441 TDALMTGYTDPDSVIDONVAVTQVDPFSLDPTFTITTOIVPODVAVSQRGRGTGRGL 1500
Db      1441 TDALMTGYTDPDSVIDONVAVTQVDPFSLDPTFTITTOIVPODVAVSQRGRGTGRGL 1500
Qy      1501 GITYYVSTGERASGMPFSVULCECYDAGAAAYELTBSETTVRLRAYNTTGLPYCQDHL 1560
Db      1501 GITYYVSSGEGSPGSMFDSVVPCECYDAGAAAYELTPAETTVRLRAYNTTGLPYCQDHL 1560
Qy      1561 FWEAVFGLTHIDAHFLSOTKQSGENFAVLTAYOATCARAKAPRPMDDMKCLTFLKP 1620
Db      1561 FWEAVFGLTHINAHFLSOTKQSGENFAVLTAYOATCARAKAPRPMDDMKCLTFLKP 1620
Qy      1621 TLVGPFTLVLRLGSVTNEVTLTHPTVKYIATCMQADLEWMTSTWLAGVLAAYVACLA 1680
Db      1621 TLVGPFTLVLRLGAVTNEVTLTHPTVKYIATCMQADLEWMTSTWLAGVLAAYVACLA 1680
Qy      1681 TGCYCITIGRLHINORAVVAPKVELYAPDEMECASAALIEBQRIAEMLKSKIOGLL 1740
Db      1681 TGCISIIIGRLHINDRVVVTPEKELIYAPDEMECASAALIEBQRIAEMLKSKIOGLL 1740
Qy      1741 QOASKOADIQPTVOASMPKVEQFMAHMMNFISIGIYLAGLSTLPQNPVAVSMMAFSA 1800
Db      1741 QOATROAQOGOPAIQSSMPKLEQFMAHMMNFISIGIYLAGLSTLPQNPVAVSMMAFSA 1800
Qy      1801 LTPSLSTSTILLINILGGLASQIAPPAGATGVFVSGVGAAYVSGISGLKVLVDIILAGY 1860
Db      1801 LTPSLPSTSTILLINIMGMLASQIAPPAGATGVFVSGVGAAYVSGISGLKVLVDIILAGY 1860
Qy      1861 AGISGALVAPKINGEKSPMEDVYNLLPGILSPALVVGVCALILRRHVGPEGAGVQM 1920
Db      1861 AGISGALVAPKIMGEEKPVEDVYNLLPALISGALVVGVCALILRRHVGPEGAGVQM 1920
Qy      1921 NRLIAFASRGNHVAPTHYVTESDASORVTOULLGSLTITSLRLAHNITEDECPICGGSW 1980
Db      1921 NRLIAFASRGNHVAPTHYVTESDASORVTOULLGSLTITSLRLAHNITEDECPICGGSW 1980
Qy      1981 LRDVWDVVCITLTDPKWMLTSKLPKMPGLPFVSCQKGYGVAGTSMITTRPCGANIS 2040
Db      1981 LQDIDMWVCISITLTDPKWMLTSKLPKMPGLPFISQKGYGVAGTSMITTRPCGANIS 2040
Qy      2041 GNVFLSGMRITGRTCKNINWGTFFPINCYTBGQVPRAPRPFKALIRVVAASEVLETQH 2100
Db      2041 GNVFLSGMRITGRTCKNINWGTFFPINCYTBGQVPRAPRPFKALIRVVAASEVLETQH 2100

```

```

Db      2041 GHWKMGTKMTGPKTCLNLMQGTFFPINCYTEGSPCVKPEPNYKTAIRVVAASEVLETQH 2100
Qy      2101 GSIVHYITGLTNDNLKVPCCOLPSPPEFFSWVDGQIHRFAPTPKPFPRDEVSFCGLNSFVY 2160
Db      2101 GSFSYVTGLTNDNLKVPCCQVPAPEFFSWVDGQIHRFAPVBPFRDEVTFTVGLNSFVY 2160
Qy      2161 GSQLPCEBPDVDVLSMLTDBSHITAEFTAARLARLARGSPSSASSASQASAPSLPATCT 2220
Db      2161 GSQLPCEBPDTEVLASMLTDBSHITAEFTAARLARLARGSPSSASSASQASAPSLPATCT 2220
Qy      2221 THGKAYDVUMDANLFMGCDVTRISGSKVYVLLDLDLPMVEERSDLEPBISEEMLPKR 2280
Db      2221 THKTAAYDCMVANLFMGCDVTRISGSKVYVLLDLDLSTTEVEDREBPSEVYLKRRK 2280
Qy      2281 PPPALPAMARPDYNPLVESWMRPYQAPATVAGCALPPRKTPTPPRRRRRTVLSDESI 2340
Db      2281 PPPALPAMARPDYNPLVETWKRPGEPFTVGCALPPLQTPVPPRRRRRAKILTDQDV 2340
Qy      2341 GVALQOLAKSFQPPSGDGLSTGAGAADSQGTPEDELASETGISSNPPLGEGIG 2400
Db      2341 EGLRBMADKVLSPLODNNDSGHSITGADTGCDIVQOPSDETAASEAGSISSNPPLGEGIG 2400
Qy      2401 DPDLPEPOYBPQPPQGVAAAPGSDSGSWSTCSBEDSVVCCSMYSWTGALITPCSPE 2460
Db      2401 DPDLPEPVGASAPSEGECEVIDSDSKSWSTVSDDESVITCCSMYSWTGALITPCSPE 2460
Qy      2461 EKLPIPLNSLSLRHYNKYCTTTSASLRAKKTFDRAQVQVDSYXSVLKDILKLAASKY 2520
Db      2461 EKLPIPLNSLSMRHYNKYCTTTSASLRAKKTFDRAQVQVDSYXSVLKDILKLAASKY 2520
Qy      2521 TALLTMEBAQOLTPPHSARKYGFQAKESVLSGAVNHIKSVKDLIEDSETPPTT 2580
Db      2521 GARLLTVEBAQALTPPHSARKYGFQAKESVLSRAVNHISVWENLIEDRFTIDTPI 2580
Qy      2581 MAKNEVFCVDPTRKGKKARLIVPDLGVRVEKALYDITQKLPALVGSYGFQYSPA 2640
Db      2581 MAKNEVFCIDPTKGGKARLIVPDLGVRVEKALYDITQKLPALVGSYGFQYSPA 2640
Qy      2641 ORVEFLKAMAEKQPMGFSYDTRCGDYTERDRIETESIVRACSLPEAHTAHSLTE 2700
Db      2641 BRVDPLKAMGSKQPMGFSYDTRCGDYTERDRIETESIVRACSLPEAHTAHSLTE 2700
Qy      2701 RLYVGGPMFNSKGQTCGRCPASGVLTSMGNTTTCYKALAAKAGIAPMLVCGD 2760
Db      2701 RLYVGGPMFNSKGQTCGRCPASGVLTSMGNTTTCYKALAAKAGIAPMLVCGD 2760
Qy      2761 DLVVISSEQCTEDERNLAPFTEAMTRVSAAPGDPREPYDELELITSCSSNVSVLGPQ 2820
Db      2761 DLVVISSEQCTEDERNLAPFTEAMTRVSAAPGDPREPYDELELITSCSSNVSVLGPQ 2820
Qy      2821 RRRVYLTDPPTPIARAAMETVRHSFVNSWLNIIQYAPTIIWABVLMTHFFSIIAODT 2880
Db      2821 RRRVYLTDPPTPIARAAMETVRHSFVNSWLNIIQYAPTIIWABVLMTHFFSIIAODT 2880
Qy      2881 LDONLNFEMYGAVYSVPLDPAIIRLHLGDLAFSLHTTYPHELTRVVASALKGAPLR 2940
Db      2881 LNONLNFEMYGAVYSVPLDPAIIRLHLGDLAFSLHTTYPHELTRVVASALKGAPLR 2940
Qy      2941 AKMSRARAVRASLISNGRBAACVGRYLFMWAVKTKLPLLEARLLDISMFTVGAAGG 3000
Db      2941 AKMSRARAVRASLISNGRBAACVGRYLFMWAVKTKLPLLEARLLDISMFTVGAAGG 3000
Qy      3001 DIHVSARAPRLLLFGLLLLFVGVGLFLLPAR 3033
Db      3001 DIHVSARAPRLLLFGLLLLFVGVGLFLLPAR 3033

```

RESULT 4
 US-09-539-601-3
 ; Sequence 3, Application US/09539601C
 ; Patent No. 6630343
 ; GENERAL INFORMATION:
 ; APPLICANT: Barteneschlager, Ralf FW

```
/ TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
/ FILE REFERENCE: all sequences
/ CURRENT APPLICATION NUMBER: US/09/539,601C
/ CURRENT FILING DATE: 2001-08-30
/ EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
/ EARLIER FILING DATE: 1999-04-03
/ NUMBER OF SEQ ID NOS: 51
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 3
/ LENGTH: 3010
/ TYPE: PRT
/ ORGANISM: Hepatitis C virus
/ US-09-539-601-3

Query Match      74.6%; Score 12027.5; DB 4; Length 3010;
Best Local Similarity 71.8%; Pred. No. 0;
Matches 2182; Conservative 352; Mismatches 470; Indels 35; Gaps 7;

Qy      1 MSTNPKPQKTKRNTNRRRPODYKFPGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
Db      1 MGTNPKPQKTKRNTNRRRPODYKFPGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60

Qy      61 RROPIPKORRSTGSKGKGPWPLYGNBGLGAGWLLSPGSRPSWGPNDPRHSRNVG 120
Db      61 RROPIPKARQPEGRAMAQGPWPLYGNBGLGAGWLLSPGSRPSWGPNDPRHSRNVG 120

Qy      121 KVIDLTGSPADLMGYIPVYGA PLGGVABALAHGVRVLEDCVNFANQNLPGCSFSIFLLA 180
Db      121 KVIDLTGSPADLMGYIPVYGA PLGGVABALAHGVRVLEDCVNFANQNLPGCSFSIFLLA 180

Qy      181 LLSCTITPVSAAEVKNIISTGVYVNTDCTNDSITWQLOAAVLHVPGCVPECKYGNASQCVI 240
Db      181 LLSCTITPVSAAEVKNIISTGVYVNTDCTNDSITWQLOAAVLHVPGCVPECKYGNASQCVI 240

Qy      241 PVSNNVAVORPGALTOGLRTHIDMVVMSATLCSALYVGDLCGGVMLAAMFIVSPQHMF 300
Db      241 ALPTLAARNASVPTTTRIRHVDLLVGAALCSAMVYGDLCGSVFLVAQLTFSPRHRBT 300

Qy      301 VODNCSTIYRGTITGHRAMDMMMMSPTATMILAYAMRVEVIDIISGAMGVWFGLA 360
Db      301 VODNCSTIYRGTITGHRAMDMMMMSPTALVLSQRLRIPOAVDMVAGAMVYLAAGLA 360

Qy      361 YFSMGAAKVVVILLAGVDARTHTVGSAAQTGRLTSLFDMGPROKIDLVNTNGSM 420
Db      361 YFSMGAAKVVVILLAGVDARTHTVGSAAQTGRLTSLFDMGPROKIDLVNTNGSM 420

Qy      421 HINRTALNCNDSLHTGFIASLPFTHS FNSSGCPBERMSACRSIEARVWGALOYEDNVN 480
Db      421 HINRTALNCNDSLHTGFIASLPFTHS FNSSGCPBERMSACRSIEARVWGALOYEDNVN 480

Qy      481 PEDMRPYCMHYPPROCGVVSACTVCGPVYCFPTSPVYVGTTRRLGAPPTTNGENETDVL 540
Db      481 PEDMRPYCMHYPPROCGVVSACTVCGPVYCFPTSPVYVGTTRRLGAPPTTNGENETDVL 540

Qy      541 INSTRPPLGSMWGCITMANSNGYTKTGCAPPCRTBADFNASTDLCPTDCFRGHPPTTYLK 600
Db      541 INSTRPPLGSMWGCITMANSNGYTKTGCAPPCRTBADFNASTDLCPTDCFRGHPPTTYLK 600

Qy      601 CGSGFWLTPRCLIDYPRILMHY PCTVNYTIFKIRMYGVGVREHRLTAACNFTRGDRCNLE 660
Db      601 CGSGFWLTPRCLIDYPRILMHY PCTVNYTIFKIRMYGVGVREHRLTAACNFTRGDRCNLE 660

Qy      661 RDRGSLSPILHSTETMALLPGSYSDLPASTGLLHONIVUDVQMYGSLPALTKIYVM 720
Db      661 RDRGSLSPILHSTETMALLPGSYSDLPASTGLLHONIVUDVQMYGSLPALTKIYVM 720

Qy      721 EWWVILFLILADARVACILMLLILGQAAALAEKLVIIILHSAASACNGFLYVIFVVAAM 780
Db      721 EWWVILFLILADARVACILMLLILGQAAALAEKLVIIILHSAASACNGFLYVIFVVAAM 780

Qy      781 YIKGAVVPLATYSLTGLMSFSLILLALPOQAAVAYDASVHGOIGALLVMTLFTLTGPKY 840
Db      781 YIKGAVVPLATYSLTGLMSFSLILLALPOQAAVAYDASVHGOIGALLVMTLFTLTGPKY 840

Qy      840 YIKGAVVPLATYSLTGLMSFSLILLALPOQAAVAYDASVHGOIGALLVMTLFTLTGPKY 840
Db      840 YIKGAVVPLATYSLTGLMSFSLILLALPOQAAVAYDASVHGOIGALLVMTLFTLTGPKY 840
```

```
Qy      841 TLLSRPLMWLCYLLTLTGEAMVOEMA PPMOVRGRDGIIMAWALIPYGVVFDITKMLAVL 900
Db      841 TLLSRPLMWLCYLLTLTGEAMVOEMA PPMOVRGRDGIIMAWALIPYGVVFDITKMLAVL 900

Qy      896 LFLARLIMWLOYFTRBAEHLQVMIPLLVNRGRDAVILLTCAIHBEFLTITKILAL 960
Db      896 LFLARLIMWLOYFTRBAEHLQVMIPLLVNRGRDAVILLTCAIHBEFLTITKILAL 960

Qy      901 GPAYLLKGLTRVPYFVRHALLRMCTMARHLAGRYYVOMALLALGRWGTYYIDHLTPM 960
Db      901 GPAYLLKGLTRVPYFVRHALLRMCTMARHLAGRYYVOMALLALGRWGTYYIDHLTPM 960

Qy      961 SPMAASGLRDLAVANVPPIIFSPMEKKVIVMGAEATAACGDIHLGLPYSAALGBVLLGPD 1020
Db      961 SPMAASGLRDLAVANVPPIIFSPMEKKVIVMGAEATAACGDIHLGLPYSAALGBVLLGPD 1020

Qy      1021 GYTSKWSLAPITTYAQTGRLGLTIYVSMGRKTEBOAGEIOYLSITVTOSPLGTSISG 1080
Db      1021 GYTSKWSLAPITTYAQTGRLGLTIYVSMGRKTEBOAGEIOYLSITVTOSPLGTSISG 1080

Qy      1077 VCMTYVYHAGSKTLAGRKPITQWYTNVDQDLVGMQAPPGASLTPCTGSSDLYLTRH 1136
Db      1077 VCMTYVYHAGSKTLAGRKPITQWYTNVDQDLVGMQAPPGASLTPCTGSSDLYLTRH 1136

Qy      1137 ADVIPARRRGRSGSILSPRPVSYLKSSGGPILCPSGHAVGI FFAAVCTRGAVAAVDV 1196
Db      1137 ADVIPARRRGRSGSILSPRPVSYLKSSGGPILCPSGHAVGI FFAAVCTRGAVAAVDV 1196

Qy      1201 PVETLDTYRSPFSDNSTPPAVPOTQVGYLHAPFGSGSKRPVAAVYAAQGVKLVNRP 1260
Db      1201 PVETLDTYRSPFSDNSTPPAVPOTQVGYLHAPFGSGSKRPVAAVYAAQGVKLVNRP 1260

Qy      1261 SVAAATLGFAGAVLSKAHGINPNI RTGRTVTTGAPITTYSGFKFLADGGCAGAYDIIICD 1320
Db      1261 SVAAATLGFAGAVLSKAHGINPNI RTGRTVTTGAPITTYSGFKFLADGGCAGAYDIIICD 1320

Qy      1321 ECHAVDSTTILIGITVLDQAEYAGVRLVYLAATPPGSYTTPHNIEBYALCOBGEIPY 1380
Db      1321 ECHAVDSTTILIGITVLDQAEYAGVRLVYLAATPPGSYTTPHNIEBYALCOBGEIPY 1380

Qy      1377 GHAIPLETIKGGRHLIFCHSKKKCDELAHKSGLDANNAAYRGADVSITPISGDVIVA 1436
Db      1377 GHAIPLETIKGGRHLIFCHSKKKCDELAHKSGLDANNAAYRGADVSITPISGDVIVA 1436

Qy      1441 TDALMTGYTGFDSVIDCNVAVTQVVDSPFTPTTITTOIVPODASVRSQRGRGTGRGL 1500
Db      1441 TDALMTGYTGFDSVIDCNVAVTQVVDSPFTPTTITTOIVPODASVRSQRGRGTGRGL 1500

Qy      1501 GIYRYVTSGERASGFPDSVILCECYDAGANWEYELTPSETTYRLRAYENTPGILPYCODHLE 1560
Db      1501 GIYRYVTSGERASGFPDSVILCECYDAGANWEYELTPSETTYRLRAYENTPGILPYCODHLE 1560

Qy      1561 FWEAVFTGLTHIDAFLSOTKOSGENFAYITAYOATVACARAKAPPSMDVMWKKCLTRLKP 1620
Db      1561 FWEAVFTGLTHIDAFLSOTKOSGENFAYITAYOATVACARAKAPPSMDVMWKKCLTRLKP 1620

Qy      1621 TLVGPPLLYRLGSTVNEVTLTHPYTKYIATCMQADLEVMSTWVLGAGVLAVAAYCIA 1680
Db      1621 TLVGPPLLYRLGSTVNEVTLTHPYTKYIATCMQADLEVMSTWVLGAGVLAVAAYCIA 1680

Qy      1677 TGSVIVIVRILISGPAIIPREVLRYREFDEMBECASHLPYIEQCMQALAEQKOKAIGLL 1736
Db      1677 TGSVIVIVRILISGPAIIPREVLRYREFDEMBECASHLPYIEQCMQALAEQKOKAIGLL 1736

Qy      1741 QOASQOADIPTVOASMPKQEPFAKMMNPFISGIQYLAGISTLPGNPAVSMMAFSA 1800
Db      1741 QOASQOADIPTVOASMPKQEPFAKMMNPFISGIQYLAGISTLPGNPAVSMMAFSA 1800

Qy      1797 QATKQAAAPAVVESKRRITLBAFAKMMNPFISGIQYLAGISTLPGNPAVSMMAFSA 1796
Db      1797 QATKQAAAPAVVESKRRITLBAFAKMMNPFISGIQYLAGISTLPGNPAVSMMAFSA 1796

Qy      1801 LTPSLSTTILNLILGMLAQALPAPGATGVVSGVGAAGVSGIGKYLVDLTIAGYG 1860
Db      1801 LTPSLSTTILNLILGMLAQALPAPGATGVVSGVGAAGVSGIGKYLVDLTIAGYG 1860

Qy      1797 ITSPLTHTTILFNILIGGMVAQAQALPASAASFVAGIAGAAGVSGIGKYLVDLTIAGYG 1856
Db      1797 ITSPLTHTTILFNILIGGMVAQAQALPASAASFVAGIAGAAGVSGIGKYLVDLTIAGYG 1856

Qy      1861 AGISGALVAFKIMSGEKSMEDVNNLPGIISPGALVGVYCAALLRHVGPBEGAVQMM 1920
Db      1861 AGISGALVAFKIMSGEKSMEDVNNLPGIISPGALVGVYCAALLRHVGPBEGAVQMM 1920

Qy      1857 AGVAGALVAFKIMSGEKSMEDVNNLPGIISPGALVGVYCAALLRHVGPBEGAVQMM 1916
Db      1857 AGVAGALVAFKIMSGEKSMEDVNNLPGIISPGALVGVYCAALLRHVGPBEGAVQMM 1916
```

Qy	1921	NRLTAFASRGHNHVAPTHYVTHESDASQRYTOLLSLTTSLRLRLHNMITECP1PCGGSW	1960
Db	1917	NRLTAFASRGHNHVSPTHYVTHESDAAARTQTLSSLTITTLKLRHOMINEDCS7PCGSW	1976
Qy	1981	LRDVMDVWCTLTLPDFKMLTSKLPPKMGPLPFVSCQKQYKGVMAGTGIMTRCCGANIS	2040
Db	1977	LRDVMDMICTVLTDFKTMLOGSKLPRLPBGVFPFCQKQYKGVMDGIMQTTCCGAQIT	2036
Qy	2041	GNVRLGSMRITGPKTCMNIWQGTPEPINCYTEGQCVKPAENFKVAILMPVAASEAYTQH	2100
Db	2037	GHWKNGSMRIVGPRPTCSNTWHTGTPPINAATTGPTCPSPAPMYSSALNRVAEEYVEYTRV	2096
Qy	2101	GSYHYITGLTNDLKNVPCOLSPPEFESWVDQVQIHRFAPTKPFPRDQVSCVLNSFV	2160
Db	2097	GDFHYVTGMTDNNKCPQVPAPEFPEVDVRLHRAAPACKPLLRBYTPELVGLNDLV	2156
Qy	2161	GSQJPCPEPPTDVLMSWLTDPSSHITAEATARLARLARGSPPEASASSASQLPAPLRACT	2220
Db	2157	GSQJPCPEPPTDVLMSWLTDPSSHITAEATARLARLARGSPPEASASSASQLPAPLRACT	2216
Qy	2221	THGAYDVWDVANI.F---MGDVTRIESGSKYVVLDSLDPMVEERSDLEPSIPSEYML	2276
Db	2217	TRHDSPADLITAEANILMQEWGNGNITRVESENKVILDSPEPLQAEDEBEREVSYPAEILR	2276
Qy	2277	PKKGFPPPLPMANRPDYNPRLVESMKRPDYQAPYAVGALPPPKTPIPPRRRRRTYGLS	2336
Db	2277	PKKGFPPPLPMANRPDYNPRLVESMKRPDYQAPYAVGALPPPKTPIPPRRRRRTYGLS	2336
Qy	2337	EDSFGDALQOLAIISFGQPPPSG--DSGLSTAGAADSCGPPPELASEGSISSMPPL	2395
Db	2337	ESTYSALAEIATYTFGSSSSAIVDSGATATA----SPDQSDDGAGSDVESISSMPPL	2391
Qy	2396	EGELGDPLDEPEQVQPPQPGQVAAAPSDSGSWTCSSE--DSDVYCCSMSSYMTGALIT	2454
Db	2392	EGEGGDPL-----SDGSMSTVSEASESDVYCCSMSTYMTGALIT	2431
Qy	2455	PCSPEBEKPLNPJLSNLSLRHNNKYCTTTSKASLRKACTVPDRMQVLDYSYVLKDIK	2514
Db	2432	PCAAEETKLPINALSNSLRHHNNLYAATTSASASLRQKQVFPDRLOVDHRYDLKEMK	2491
Qy	2515	LAASKTARLLTMBEACOLTPEPHASRYXGGAKEVRSLSGRAVNHIKSWKDLLESET	2574
Db	2492	AKASTVAKKLSVEBACKLTPPHSRRSFKFGAGADVRLSSKAVNHIRSWKDLLEDET	2551
Qy	2575	PIPTTIKAKNEVFCVDPYTGKGAARLLVYVYDLAGRVEKALYDITOKLPOAVMGASYG	2634
Db	2552	PIDTTIAKAKNEVFCVQPEKGRKRPARLLVFDFLAGRVEKALYDVYSTLPQAVMGSSYG	2611
Qy	2635	FQYSPADQVFEFLKAMAEEKDPMGESYDTRCPDSTVTERDITREBSIYRACSLPEBAHTA	2694
Db	2612	FQYSPADQVFEFLVAMAKKCCPMGFAAYDTRCFDSTVTEINDIRVESIYQCCDLPEARQA	2671
Qy	2695	IHSILTERLYVGGPMFNSKQTCGVRGCAAGVLTSMKNTITTCYKALAAKAGIAPT	2754
Db	2672	IRSLTERLYIGPPLNRSKQNGCGRRCASGVLTSSCNTLTTCYLKAAAKARAKLDOCT	2731
Qy	2755	MLVCGDDLVVYSESQGTEDERNRLNFAFEATMTRYSAAPGDPPREYDELITSCSNVS	2814
Db	2732	MLVCGDDLVVYCESAGTQEDBASLRFAFEATMTRISAPGDDPKPEYDELLITSCSNVS	2791
Qy	2815	ALGFGRRRYVLTDPPTPIARAAMETVRHSPVNSWLNIIQYAPTIWARMVLTMPFSI	2874
Db	2792	AHDASGKKVYVLTDPPTPLARAAMETARHTPVNSWLNIIIMYAPTLLWARMVLTMPFSI	2851
Qy	2875	LMADDTLDQNLNFMFYGAIVSVSPDLPAIIEERLHGLDAFSLHTYTPHELTRVASALRKL	2934
Db	2852	LLAEOLEKALDQCQYGCYSIEPDLPOIILQRLHGLSAFSLHSYSPBEINRVASCRLK	2911
Qy	2935	GAPPLRAKSRARARASLISRGGAACVGRILPFWMAKTYKLKTLPLPEALTLDDSSWFT	2994
Db	2912	GVPPRLVRHRRARSRARALLSQGGRAACGKTLFWMAKTYKLKTLPLPAASQDLSSWFT	2971
Qy	2995	VGAGGGDLYHSVSRAPRLDLFGLLLFVGVGLFLPLPAR	3033

DB	29712	AGYSGGDYHSLSRARRPWFMMCLLLSVGVGYLLFLNR	3010
		<p>RESULT 5</p> <p>US-09-539-601-21</p> <p>; Sequence 21, Application US/09593601C</p> <p>; Patent No. 6630343</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: Batecheslager, Ralf FW</p> <p>; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System</p> <p>; FILE REFERENCE: all sequences</p> <p>; CURRENT APPLICATION NUMBER: US/09/539,601C</p> <p>; CURRENT FILING DATE: 2001-08-30</p> <p>; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY</p> <p>; EARLIER FILING DATE: 1999-04-03</p> <p>; NUMBER OF SEQ ID NOS: 51</p> <p>; SOFTWARE: Patentlin Ver. 2.1</p> <p>; SEQ ID NO: 21</p> <p>; LENGTH: 3010</p> <p>; TYPE: PRT</p> <p>; ORGANISM: Hepatitis C virus</p> <p>US-09-539-601-21</p>	
QY		Query Match	74.5%; Score 12010.5; DB 4; Length 3010;
		Best Local Similarity	71.7%; Pred. No. 0;
		Matches 2178; Conservative 355; Mismatches 471; Indels 35; Gaps 7;	
DB	1	MSINPKPORKTKRNTNRRODVKKPFGGQIYGVYLLRRRGRGLRVARTKTSERSOPRG	60
QY	61	RRQPIPKRRRSTGKSGKPGYWPWLYNGEGLGMAGMLTSPGSRPSPGPNDRHRSRVNG	120
DB	1	MGINPKPORKTKRNTNRRODVKKPFGGQIYGVYLLRRRGRGLRVARTKTSERSOPRG	60
QY	61	RRQPIPKRRRSTGKSGKPGYWPWLYNGEGLGMAGMLTSPGSRPSPGPNDRHRSRVNG	120
DB	61	RRQPIPKRRRSTGKSGKPGYWPWLYNGEGLGMAGMLTSPGSRPSPGPNDRHRSRVNG	120
QY	121	KVIDTLTTCGPAIDMGYIPVVGAPLGGVARALAHGVRVLEDGYNATGMLPGCSFSIFLLA	180
DB	121	KVIDTLTTCGPAIDMGYIPVVGAPLGGVARALAHGVRVLEDGYNATGMLPGCSFSIFLLA	180
QY	181	LISCITTPVSAAEVKNISTGYMTNDCNDSITWGLAAVLHPVPCVCEKYNASOCMI	240
DB	181	LISCITTPVSAAEVKNISTGYMTNDCNDSITWGLAAVLHPVPCVCEKYNASOCMI	240
QY	241	PVSPVAVAVQRGALTOGLRTHIDMVMSATLCSALYVGDLCGGWLAQAQMFVSPQHWF	300
DB	241	ALTPFLAARNASVPTTTRRRHVDLLVGAALCSAMTVEDLCSGVFLVAQLFFFSRPHET	300
QY	301	VQDNCNSIYPGTITGHRMADMMNMSPTATWILAYAMRVBEVILDIISGAHWGMFGIA	360
DB	301	VQDNCNSIYPGHVTGHRMADMMNMSPTALIVYSQLIRIPQAVVDVMAAGHWGVLAGIA	360
QY	361	YFSNOGAMAKVIVILLAAGYDARPHYVGGSAQOTGSLTSLFDMGPRROKLOLVNTGNSW	420
DB	361	YYSWGNNAKXLYIWLFLPAGVDGGYTVVGGTMAKNTLITSLTFSFGSSQKLOLVNTGNSW	420
QY	421	HINRTALNCNLSLHTGFASLFTYTHSPNNSGCEPESMASCRSIEAFRVGMALOYEDNYTN	480
DB	421	HINRTALNCNDSLNTGFLAALFYHKFNSSGCEPESMASCRSPIDAFAGMGPTTY--NESH	478
QY	481	PEDMRPYCMHYPPROCGVAVSAKTVCGPYVCTPSPVVVGTTRDLGAPTYTWGENETDVFLL	540
DB	479	SSDQRPYCMHYAPRCGIVPAAYOVCGPYVCTPSPVVVGTTRDFGVPTYSWGENETDVL	538
QY	541	LNSTRPPLPGSNFGCTMMNNSGTYTKCGAPPCRTADFNASIDLCTPTDCFRHHPTTYLK	600
DB	539	LNSTRPPOGNWFGCTMMNNSGTYTKCGAPPCNIGIGIKKT--LTCPTDCCFRHHPATYTK	596
QY	601	CGSGGWLMPRCGLIDVPRYLMHYPCVANTTIFIKIRMYVGVGEHRLTAAACNFTGDCNCLBD	660
DB	597	CGSGGWLMPRCGLVHYPRYLMHYPCVANTTIFIKIRMYVGVGEHRLTAAACNFTGDCNCLBD	656
QY	661	RDRGSLSLHSTTEWAILPCSYSDLPALSLGTLAHQNIADVQEMYGSLSPALTRYIVRW	720

Db 657 RDRSELSPLSTTEWVLPSCFPTLPALSTGLIHLQNVVDQVYLXIGISAVVSPAIXW 716
Qy 721 EWTLLPLLLADNRVACLMMLLLGQAEALFKVYLTHAASAASCGFLYPIFFPAAM 780
Db 717 EYVLLPLLLADNRVACLMMLLLGQAEALFKVYLTHAASAASCGFLYPIFFPAAM 776
Qy 781 YIKRVPPLATYSLTGMSFSLLLALPOQAVYADASVHQIQAALLMVLITLTPRGYK 840
Db 777 YIKRVLPGAAVYALXGWPPLLLLLALPPRAYAMDREMAASCGAVFVGLILITLSPHYK 836
Qy 841 TLLSPLMMLCYLLTLEAMVQEMAPMOVGRGDIIMAVAFYRQGVEDITKMLLAVL 900
Db 837 LFLARLIMLQYFTLBAEHLQVMIPLNVRGRDAVILLTCAIHPELLFTIKILLAL 896
Qy 901 GPAYLKGALTRPRYPYRABALLRMCTMAHLAGKRYVOMALLAGMVTGTYYDHLTPM 960
Db 897 GPLMVLQAGITTKVPYFRAHGLIRACMLVRKAVGHVOMALMKLALTGTYYDHLTPL 956
Qy 961 SDMAAGRLDAVAVEPIFSPMEKKYIWMGAETAAAGDILHGLPVSARLGRREVLPAD 1020
Db 957 RDMHAGLNDLAVAVEVPFSDMETKVTITWGAADTAACGDIILGLPVSARLGRREVLPAD 1016
Qy 1021 GYTSKMSLLAPITAYAOQTRGLLGTIVVSMTRDKTEBAGEIQVLTSTVQSFAGTSISG 1080
Db 1017 SLEGGCMRLAPITAYSOQTRGLLGTITSJLGRDRNQVEGVYVSTAQSFATCVNG 1076
Qy 1081 VLMTVYHAGNKTLASSRGPVTOMYSSABDLYGMSPPECTKSLECTGCANDLYLVTRN 1140
Db 1077 VCMTVYHAGSKTLAPKGPITQMYTNVODLVMQAPGARSLTCTCGSSDLVLYVTRH 1136
Qy 1141 ADVIPARRRGDKKALLSPPLSTLKSSSGGPLYCRGHAUVFRAVACSRAVAKSIDP 1200
Db 1137 ADVIPARRRGDKKALLSPPLSTLKSSSGGPLYCRGHAUVFRAVACSRAVAKSIDP 1196
Qy 1201 PVETLIDIVTRSPFSDNSTPPAVPQTVQVGYLHAPTSGSKSTVPVAYAAQYKVLVLP 1260
Db 1197 PVSMETMTRSPFVTDNSPPAVPQTFQVYHLLHAPTSGSKSTVPVAYAAQYKVLVLP 1256
Qy 1261 SVAAITGFGAYLSKANGINPITRTYRTYTGAPRTYTYGKRLADGGCAGAYDIIICD 1320
Db 1257 SVAAITGFGAYLSKANGINPITRTYRTYTGAPRTYTYGKRLADGGCAGAYDIIICD 1316
Qy 1321 ECHAVDSTIIGITGLDOAETAGVRLTVLATATPPGSVTPPHNIEVVALQGEGRIPY 1380
Db 1317 ECHAVDSTIIGITGLDOAETAGVRLTVLATATPPGSVTPPHNIEVVALQGEGRIPY 1376
Qy 1381 GRAIPLSYIKGGRHLIFCHSKKKCDELAALRGMGINSVAYYRGLDVSYIPTOGDVVVA 1440
Db 1377 GRAIPLATIKGGRHLIFCHSKKKCDELAALRGMGINSVAYYRGLDVSYIPTOGDVVVA 1436
Qy 1441 TDALMTGYGDPSPVSDCNVAVTQVDFSLDPTFTTTOIVPODAVSRSQRRGRTGRGL 1500
Db 1437 TDALMTGYGDPSPVSDCNVAVTQVDFSLDPTFTTETTTVPDAVSRSQRRGRTGRGM 1496
Qy 1501 GIYRYVSTEBASGMDSVYLCECYDAGAAWELTSETTVLRRAVFNTPGLPVCODHLE 1560
Db 1497 GIYRFYTPGBRSGMDSVYLCECYDAGAAWELTSETTVLRRAVFNTPGLPVCODHLE 1556
Qy 1561 FMEAVFTGLTHIDAHFLSQTKSGENFAYLTAQATVCARAKAPPPSPVMVMKCLTRLKP 1620
Db 1557 FMEAVFTGLTHIDAHFLSQTKSGENFAYLTAQATVCARAKAPPPSPVMVMKCLTRLKP 1616
Qy 1621 TLVGPPLLYRLGASVTNEVTLHPVTKYLATCQADLEVMTSTWVLAGVLAANAAYCLA 1680
Db 1617 TLHGPPPLLYRLGASVTNEVTLHPVTKYLATCQADLEVMTSTWVLAGVLAANAAYCLA 1676
Qy 1681 TGCVCITIGLHINORAVVAPDKVEVLEAPDEMECASRALIJBEGORIAMKSKXIGLL 1740
Db 1677 TGSVAVTVGHILILSGKRAIIPDRVLYRDEMECASRLPIYBQGNQLEBOFQKXIGLL 1736
Qy 1741 QOASKOADIQPTVQASWPKVEQFMAKHMNFIISGIQYLAGLSTLPGNPAVAVSMAFSA 1800

Db 1737 QYATQOAEAAAPVSESKWRTLEAFMAKHMNFIISGIQYLAGLSTLPGNPAVAVSMAFSA 1796
Qy 1801 LTPSPLSTIITLNLIGWLLASQIAPPAGATGVVSGVGAANVSGIGLKVNVLDIAGV 1860
Db 1797 ITSPLTQHTLPLNLIGWVVAQIAPPASASAVGVGIGAAVGSIGLKVNVLDIAGV 1856
Qy 1861 AGISGALVAFKIMSCEKSPMEDVNLPLGILSPGALVGVICAAILRRHYGPEGAVQNM 1920
Db 1857 AGVAGALVAFKIMSCEKSPMEDVNLPLGILSPGALVGVICAAILRRHYGPEGAVQNM 1916
Qy 1921 NRIAPASGNHVAPTHVYTESDASQVTOILLGSLTITSRLRLHNVITEDCPIPCGGSV 1980
Db 1917 NRIAPASGNHVSPTHVYTESDASQVTOILLGSLTITSRLRLHNVITEDCPIPCGGSV 1976
Qy 1981 LBDVMDWCTIITLDFEKNMLTSLFPMRGLPVSOCKGKYGWAGTGIMTTCCPGANIS 2040
Db 1977 LBDVMDWCTIITLDFEKNMLTSLFPMRGLPVSOCKGKYGWAGTGIMTTCCPGANIS 2036
Qy 2041 GNVRLGSMRITGPCTMINIQCCTPINCYTEGCQVPAKPNFKVAIWRVAASEYAEVTOH 2100
Db 2037 GHVKGSMRITGPCTMINIQCCTPINCYTEGCQVPAKPNFKVAIWRVAASEYAEVTOH 2096
Qy 2101 GSYHYITGLTJTNLKVPCQLPSPEFSSVVDVQIHRFADTPKPFREDEVSFCVGLNSFV 2160
Db 2097 GDFHYITGLTJTNLKVPCQLPSPEFSSVVDVQIHRFADTPKPFREDEVSFCVGLNSFV 2156
Qy 2161 GSOLPCDPEPBDVYLSMKTLDPSHITAETARLARGSPSESSASSQSLASPLAATCT 2220
Db 2157 GSOLPCDPEPBDVYLSMKTLDPSHITAETARLARGSPSESSASSQSLASPLAATCT 2216
Qy 2221 THGKADVMDVANLF---MGADVTRIESGKVVVLDLDPMEBERSLSPSIPSEYML 2276
Db 2217 THGKADVMDVANLF---MGADVTRIESGKVVVLDLDPMEBERSLSPSIPSEYML 2276
Qy 2277 PKRREPPALPAMARPDVNPPLVESWKRPDYOPATVAGCALPPRKRTPPERRRTVGLS 2336
Db 2277 PKRREPPALPAMARPDVNPPLVESWKRPDYOPATVAGCALPPRKRTPPERRRTVGLS 2336
Qy 2337 EDSIGDALQOLAIKSPGQPPSG--DSGLSTGGAADSGQTPPDELALSETGISISMPL 2395
Db 2337 EDSIGDALQOLAIKSPGQPPSG--DSGLSTGGAADSGQTPPDELALSETGISISMPL 2391
Qy 2396 EGEGLDPEPEQVPEPPQCGVAAFGSDSGSWTSCSE--DSSVVCSSMSYMTGALT 2454
Db 2392 EGEGLDPEPEQVPEPPQCGVAAFGSDSGSWTSCSE--DSSVVCSSMSYMTGALT 2431
Qy 2455 PCSPEBEKLPINPLSNSILRYHNKYCTTKSASLRKAKVTFDRMQVLDYSYDVLKDIK 2514
Db 2432 PCAAEBTKLPVVALSNSILRHHNLVYATTSRSASLRKAKVTFDRMQVLDYSYDVLKDIK 2491
Qy 2515 LMAKVTARALLMEARCOLTPHSHARSKYGFAKSEVRSLSGAANHIKSVKMDLLEDST 2574
Db 2492 AKASVTAKALSVESBECKLTPHSHARSKYGFAKSEVRSLSGAANHIKSVKMDLLEDST 2551
Qy 2575 PIPTIIMAKNEVFCVDPTEKGGKARLIYVPLGVRVCEKMLYDITQCLPOAVMGASYG 2634
Db 2552 PIPTIIMAKNEVFCVDPTEKGGKARLIYVPLGVRVCEKMLYDITQCLPOAVMGASYG 2611
Qy 2635 FOYSPAORVEFLKAMAEKDDMGFSYDTRCDSTVTERDIRTEBSIYRACSLPEBAHTA 2694
Db 2612 FOYSPAORVEFLKAMAEKDDMGFSYDTRCDSTVTERDIRTEBSIYRACSLPEBAHTA 2671
Qy 2695 IHSILTEBLYVGGPMRNSKQCTGYRCRASGVLTSMGNTTICYYKALAACGAAGIATP 2754
Db 2672 IHSILTEBLYVGGPMRNSKQCTGYRCRASGVLTSMGNTTICYYKALAACGAAGIATP 2731
Qy 2755 MLVCGDDLVVISESQCTEBEENLRAFTEAMRYAPPCDDPRPREDELTILSCSSNVSV 2814
Db 2732 MLVCGDDLVVISESQCTEBEENLRAFTEAMRYAPPCDDPRPREDELTILSCSSNVSV 2791
Qy 2815 ALGPOGRRRYLTRDPTPIABAWEYVHSPVNSWMLGNIIOYAPTIVARVYLMTHFESI 2874
Db 2792 AHDASGRVYVYLTBDPTPIABAWEYVHSPVNSWMLGNIIOYAPTIVARVYLMTHFESI 2851

[illegible]

QY	541	INSTRPPLSGMSECCWMMSSSGYTKCCGAPPCRTADFNASDIDLCPCTOCFRGHPTLYX	600
Db	539	LNMRPPOGMNPGCWMNMSGTGTCCGPPCNIIGIGINKT--LTCPTOCFRGHPRATYTK	596
QY	601	CGSGFWLTPRCLIDVPRYLMMYPCTVANTIEKIRMYVGVGEBHRLTAACNFTGRDCNLED	660
Db	597	CGSGFWLTPRCLMHPYRLMMYPCTVNTIFKVRMYVGVGEBHRLBAACNMTGRGCNLED	656
QY	661	RDRSOLSPULSHSTTEMAILPCSYSDLPALSTGLLHONVYDVQEMYGSLPALTKYRW	720
Db	657	RDRSELSPLLSTTEMOVLPCSFPTLPALSTGLLHONVVDVQYLYGIGSVAFPAIKW	716
QY	721	EMVILLFLLDADARCACTMMLLILGQEAALKEKVIHMAASAACNGEYVIFEFVAM	780
Db	717	EYVLLFLLDADARVCACCTMMLLILQEAALKEKVIHMAASAACNGEYVIFEFVAM	776
QY	781	YIKGVEVLATYSLTGLWGSFSLILLALPOQAAVAYDASHVGOIGALLWITLFTLTPGX	840
Db	777	YIKGVLVGAAYALYGVMPRLILLILLALPRPVAANDREMAASCGGAVFGLILLTLSPHX	836
QY	841	TLLSRFLMWCYLLTLGEMAOEMAPMQVRGGRDGIIIMAAVFPVGVFDITKKMLAVL	900
Db	837	LFARLIWMLOYPFTRAEAHLOWMPRLPNVNGRGDAVILLTCALHPELIFITTKILLAIL	896
QY	901	GPAYLLKCALTRVPEFVAGHALLMCTMAHLAGROYOMLTLAGRMGTIYHILPM	960
Db	897	GPLMLOQGITKVPFVRAHGLILRCMLVRKVAAGHYOMALMKLALTGIVYHILPL	956
QY	961	SDMAASGLRDVAVAEPIIFSPMEKKVIMWGAETACGDILHGLEPVSARLGREVLLGPAD	1020
Db	957	RDMAHAGLRDLAVAEVPEVFSDMETKVIWGDATACGDIIILGLEPVSARGREHILGPAD	1012
QY	1021	GYSKGSMLAPITAYAOOTRGLGTIVSMTRGRKTEOAGEHIOVLSTYTGOSFGLTSISG	1080
Db	1017	SLEGGMRLLAPITAYSOOTRGLGCITISLGRBRNOVEGQVQVSTATOSFLATCVNG	1076
QY	1081	VLMPTYHAGKTKTAGSGRPYOMWSSABGDIWGPSPGRTSLPEPCGANDLVLNRM	1144
Db	1077	VCMTYHAGSGTKTAGSGRPITOMITNVODJWQGAOPPGARSILPCTCGSDLVLTRH	1133
QY	1141	ADVIPARRRRGKRGALLSPRLSTLKGSSGGPVLCPRGHAGVFPRAAVCSRGAVSIDFI	1200
Db	1137	ADVIPVRRRRGDSRGSLSPRPVSLKSSGGPVLCPSGHAGVIFPAAVCTRGVAKAVDPV	1199
QY	1201	PYETLIDITRSPTESDNSTPPRAVPOTYOVCYIHAFTGSGSKTKVYVAAOQYKVLVLP	1266
Db	1197	PVESMETTRMSVPFTDNSSPPRAVPOTFQVAHLHAPTGSGSKTKVPAAYAAOQYKVLVLP	1255
QY	1261	SVAATLGFAGVYSKXHGINPNIIRTVRVTTCGAPITYSYGFELADGGCAGAGAYIICD	1332
Db	1257	SVAATLGFAGVYSKXHGIDPNIRIGVRITITTCGAPITISTYGFELADGGCGSAGAYIICD	1312
QY	1321	ECHAVDSTTIIIGITVLDOAETAGVRLTVLATATPPGSVTPPHNIEEVALGOEJEIPY	1380
Db	1317	ECHSDSTTIIIGITVLDOAETAGARLVVLATATPPGSVTPPHNIEEVALSSTGEIPY	1376
QY	1381	GRAPLSTYIKGGRHLIFGSKKKCBELAAALRGMLNSVAYYRGLDVSTIPQSGVVVYA	1444
Db	1377	GKAIPLEIKGGRHLIFGSKKKCBELAAKLSGLGLNVAAYYRGADVSTIPRSGVIVYA	1433
QY	1441	TDALMTGTGDPDSYIDCNVAATQVVDPSLDPFTFIITQIYVQDAAVRSORRGRTGRGL	1500
Db	1437	TDALMTGTGDPDSYIDCNTCVTQVVDPSLDPFTFIETTYIPQDAAVRSORRGRTGRGM	1499
QY	1501	GIYRYVSTGERASGMFDSVILCECYDAGAAWEILTPSETTVRLRAYFNTPLPVCOHLE	1566
Db	1497	GIYRVVTGERSGMFDSSVILCECYDAGAWELTPAETSVRLRAYLNTPLPGVCOHLE	1555
QY	1561	FWEANFTGLTHIDAHFISQTKOSGENFAYLTRYOATVCARAKAAPPSMDVMKKCLTRLP	1622
Db	1557	FWEASFTGLTHIDAHFISQTKOAGDNFPYLVAYOATVCARAOAPPSSDWMKKCLIRLP	1611

QY	1641	PLVGEPTLLYBYLGSYTNENVTLTHPTXKLIANCQADLEWMSWTWYLAGCVLAAYVAGLA	1684
Db	1617	TLHGFTPLLYKGAQVENVTTTHPTKXLIAMCSADLEVNVTWLVGCVLAALAAVCLT	1676
QY	1681	TGCVCIIGRLHINCRVAVAPDKVEVLYEAFDEMECASPAALEEGQRIAEMLKSKIQGL	1740
Db	1677	TGSVVIIVERIIILSKRPALIPPREVLYRFEDEMECASHLPYIEGQMGLAEQFKQALIDL	1733
QY	1741	QOASQOAOADIOPTVOASWPVKEQFMAKIMWNFISGIQYLAGLSTLPGNPVAVASMPAFSA	1800
Db	1737	QATKQAEAAAAPVESKRRTEAEFAKIMMNFISGIQYLAGLSTLPGNPALIASLMAFPAS	1796
QY	1801	LTSPLSTSTTLLNLGGMTLASQIAPRAGAGFVVSGLVGAAGSIGIGKYLVDLIAQY	1866
Db	1797	ITSPLLTHTHTLLFNLLGGMVAAQLAPRASAFAVAGIAGAAGSIGIGKYLVDLIAQY	1856
QY	1861	AGISGALVAFKIMSGEKSPMEDVWLLPGIISPGALVGVYCAALIRRHVGESEAQOM	1922
Db	1857	AGVAGALVAFKVMSEMPSTEDLVULLPAILSPALVGVGVCAALIRRHVGESEAVOM	1916
QY	1921	NRLIAFASRGNHVAETHYVTESDASQRYTOLLGSLTITLSLRLHNMITBEDCP1PCGGSW	1986
Db	1917	NRLIAFASRGNHVSPTHYVPESDAARVTQILSSLTITQLLKRHLQWINEDCSTPCGSGW	1976
QY	1961	LRDVMDVCTIILDPKXMYLTKLFPKMGCLPVSQCKYKGVWAGTGIWTRCPGAGNIS	2044
Db	1977	LRDVMDVCTIYLTPFKMYLTKLPRLGVPFPFSCORGYKGVWRDGIWOTTCPCGAQIT	2038
QY	2041	GNVRLGSRIRIGPKTCMNIWGSTPFINCYTEBOCVKRPAPPKVATMVAASEVAYEYQ	2100
Db	2037	GHVKKGSRIYGPFRICSTWHTGTFPINAVTTGPTCPSPAPITSRALMVAABEYEVETRV	2098
QY	2101	GSYHYITGLTJTDNLKVPQCLPSPEEFSWVDGVQIHRFAPTRKPFPRDEVSFCVGLNSFV	2166
Db	2097	GDFHYVTGMTTDDVYKCPQVAPABEFTEVDGRLHRYAPACKPLIREVTVLGVANGVLY	2156
QY	2161	GSQLPCEBEPDVTLYMSMLTDPSSHITAAARLARLARGPPESEASSASQLSAPSLRATCT	2222
Db	2157	GSQLPCEBEPDVAVYLTSMMLTDPSSHITAAARLARLARGPPPLASSASQLSAPSLKATCT	2216
QY	2221	THGKAYVDVMDANLF-----NGQDVTRIESGSKVVLVLDSPMVEERSDLSPSEMYL	2276
Db	2217	TRHDSPPDADLLEANLIMRQEMGANTRVESEKVVILDSFELOAEDEDEREVSAPAEILR	2276
QY	2277	PKKRPFPALPAMABPDYNPPLVESWKRBDYOPATVAGCALPPPKTPTPPRRRRRTVGLS	2336
Db	2277	RSKRFPFRAMPYIWARBDYNPPLLESKODDYPRVPHGCLPPLAKAPPIPRRRRRRTVGLS	2336
QY	2337	EDSIGDALIQUAISKFCQPPPSG-DGSLSTGAGAADSQTPREBALSETGISISMPPL	2395
Db	2337	ESTVSSALAEIATEFTEFGSESSAVIDSGTATA-----SPDQSDODGADSDVESYSMPPL	2390
QY	2396	EGELGDPPLBEQVQVPEPPPOGGVAAPESSDGSWSTCEEF-DDSVYCCSMYSWTGALIT	2456
Db	2392	EGBEPDPL-----SDGSMYSVEASESDVVCSCSMYWTGALIT	2433
QY	2455	PCSEPEEBECLPINPLSNSILRYHNKYCTTTSASASIRAKYTFPDRQOVDSZYDSYTLKDIK	2514
Db	2432	PCAAEETCLPINALSNSILRRHNLVYATTSASASIRKOKYTFDRQVLDHYRVLKEMK	2491
QY	2515	LAASKVTARLLIMBEACQLTPPHSARSKYFGQAEVRSLSGPAVNHISKVWKLDLDESET	2574
Db	2492	AKASVTAKALLSVEACQLTTPHSARSKYFGAGADVRMLISKAVNHIRSWKLDLEDRET	2551
QY	2575	PIPTTIMAKNVEFCVDPTRKGGKKAARLVYDPLIGRVVEKMAVLDIOTKLPOAVWAGASYG	2634
Db	2552	PIPTTIMAKNVEFCQPEKGGKKAARLVYDPLIGRVVEKMAVLDVSTLPOAVWAGSSYG	2611
QY	2635	FOYSPAQVVEFLTKMAWEKQPMGFSYDTRCFDSTVTERDRIITREESIYRASLPEEATA	2694
Db	2612	FOYSPGQVVEFLVNAWMKAKCQPMGAYDTRCFDSTVTEINDIRVESIYQCCDLAPAEARQA	2671
QY	2695	IHSILTERLYVGGPMFNISKQOTGYRRCRASGVLTSMGNITTCYVKALAAKAAGIAPT	2754

Db	2672	IRSLTEKLYIGGFLTNKSCGNGYRRRCASGVLTTSCGNILTYLKAAMAAACRAKADODCT	2731
Qy	2755	MLVCGDDLVIYSESQTEEDERNLRAFEAMTRYAPGDPPEBYDLLEITSCSNVS	2814
Db	2732	MLVCGDDLVIYCESAGTQEDBASIRAFTEAMTRYAPGDPPEKPEYDLLEITSCSNVS	2791
Qy	2815	ALGQGGRRRYLTDPPTPLARAMEVRSVNSWIGNIIOVPTIWARVMTHTPFI	2874
Db	2792	AHDASGRVYLYLTDPPTPLARAAMEVRAHRTPVNSWIGNIIMAPTLMARVIMTHTPFI	2851
Qy	2875	LMADPTLDQNLNFEEMYGAVYVSPLDPAIIEIRLHGDAESLHTYDPHELTRVASAIRKL	2934
Db	2852	ILHAEQLEKALDCQIYGACYSIEPLDIPQIIQRHGLSATSLSHSYSGEINRVASCLRKL	2911
Qy	2935	GAPPLRWKSRBARAVRASLISRGRAAVCGRYLFNMAVTKLKLTPLPEARLLDLSWFT	2994
Db	2912	GVPEPLRWHRARSRVARLLISQGGRAATCKYLFNMAVTRTKLKTPIPAASQDLDSWFW	2971
Qy	2995	VGAGGDIYHSVSRAPRLLLFGILLFVVGCLFLLPAR	3033
Db	2972	AGYSGDIYHSLSRAPRPMFMCKLLISVGGLYLLFNR	3010

US-08-324-977-2
: Sequence 2, Application US/08324977
: Patent No. 5747339
: GENERAL INFORMATION:
: APPLICANT: OKAYAMA, Hiroto
: APPLICANT: FUKE, Isao
: APPLICANT: MORI, Chiasto
: APPLICANT: TAKAMIZAWA, Akahisa
: APPLICANT: YOSHIDA, Iwao
: TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
: TITLE OF INVENTION: GNA AND ANTIGEN POLYPEPTIDE
: NUMBER OF SEQUENCES: 50
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Armstrong, Westerman, Hattori, Mcleland &
: ADDRESS: Naughton
: STREET: 1725 K St. N.W. Suite 1000
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20006
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.5 in., 1.44MB
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
: SOFTWARE: ASCII
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/324,977
: FILING DATE: 18-OCT-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 2-167466
: FILING DATE: 25-JUN-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 2-230921
: FILING DATE: 31-AUG-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 2-305605
: FILING DATE: 09-NOV-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/099,706
: FILING DATE: 30-JUL-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/769,996
: FILING DATE: 02-OCT-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/635,451
: FILING DATE: 28-DEC-1990
: ATTORNEY/AGENT INFORMATION:
: NAME: Stevens-Smith, Theresa M.

REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703D
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3010 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-324-977-2

Query Match 74.3%; Score 11984.5; DB 1; Length 3010;
Best Local Similarity 71.5%; Pred. No. 0;
Matches 2175; Conservative 352; Mismatches 472; Indels 45; Gaps 8;

QY 1 MSTNPKQQRKTNTNRRPODVKPPGGQIVGVYLLPRGRPLGVATKTSERSQPRG 60
DB 1 MSTNPKQQRKTNTNRRPODVKPPGGQIVGVYLLPRGRPLGVATKTSERSQPRG 60
QY 61 RROGIPRDRSSTGSMKPGPMPLYGNBGLGNAGMILLSPRSGPSWGPMDPHRSRNVG 120
DB 61 RROGIPRDRSSTGSMKPGPMPLYGNBGLGNAGMILLSPRSGPSWGPMDPHRSRNVG 120
QY 121 KVIDTLTCGFADLMGYIPVGAAPLGVARALAHGVRVLEDSGVNFAIGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVGAAPLGVARALAHGVRVLEDSGVNFAIGNLPGCSFSIFLLA 180
QY 121 KVIDTLTCGFADLMGYIPVGAAPLGVARALAHGVRVLEDSGVNFAIGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVGAAPLGVARALAHGVRVLEDSGVNFAIGNLPGCSFSIFLLA 180
QY 181 LLSCTTPVSAAEKNIISTGMVTNDCTNDSITWLOAAVLHPGCPYCEKRVGNASQWT 240
DB 181 LLSCTTPVSAAEKNIISTGMVTNDCTNDSITWLOAAVLHPGCPYCEKRVGNASQWT 240
QY 241 PVSNVAVVORRGALTOGRTIHDVMSATLCSALVYGDLCGGWMLAQMFIYSPOHMF 300
DB 241 PVSNVAVVORRGALTOGRTIHDVMSATLCSALVYGDLCGGWMLAQMFIYSPOHMF 300
QY 301 VQDNCISYPTGITIGHMAMDMNMNMSPTATMLAYAMRVEVILLIISGAHMGVEGLA 360
DB 301 VQDNCISYPTGITIGHMAMDMNMNMSPTATMLAYAMRVEVILLIISGAHMGVEGLA 360
QY 361 YFSQGMAMAKVYLLLAAGYDARTHTVGSAAOTGRLLTSLFPMGRKQIQLVNTNGSW 420
DB 361 YFSQGMAMAKVYLLLAAGYDARTHTVGSAAOTGRLLTSLFPMGRKQIQLVNTNGSW 420
QY 421 HINTALNCNDSLHTGFASLFYTHSFNSGCEPERMSACRSIAFRVGMALQYEDVNVN 480
DB 421 HINTALNCNDSLHTGFASLFYTHSFNSGCEPERMSACRSIAFRVGMALQYEDVNVN 480
QY 481 PEDMRPYCMHYPPROCGVNSAKTVCGPVYCTPSPVVVGTDDLGAFTYTGENSEIDVFL 540
DB 481 PEDMRPYCMHYPPROCGVNSAKTVCGPVYCTPSPVVVGTDDLGAFTYTGENSEIDVFL 540
QY 541 LNSRPRPGSWFGCTMANSSTYTCGAPPCRTTRADTNASTDLICPTDCRKRHPDTTYLK 600
DB 541 LNSRPRPGSWFGCTMANSSTYTCGAPPCRTTRADTNASTDLICPTDCRKRHPDTTYLK 600
QY 601 CGSGPMTLPRCLIDYPRILMHYPCYVNTYTKIKIRMYGVGVENHRLTACNTRGDRCLD 660
DB 601 CGSGPMTLPRCLIDYPRILMHYPCYVNTYTKIKIRMYGVGVENHRLTACNTRGDRCLD 660
QY 661 RDRSOLPLHSTHTEWALPCSYSDLPALSTGLHLHQNIVDYQFMYGLSPALTKYLVN 720
DB 661 RDRSOLPLHSTHTEWALPCSYSDLPALSTGLHLHQNIVDYQFMYGLSPALTKYLVN 720
QY 721 EFWLLPLLLADAVYCACTMMLILGQAEALAEKLVILHAASAASCNGLYFVYFPAAN 780
DB 721 EFWLLPLLLADAVYCACTMMLILGQAEALAEKLVILHAASAASCNGLYFVYFPAAN 780
QY 781 YIKRVVPLATYSYGLMSFSLILLALPQOAYAVADASHGQIGALLVMTLFTLPFGYK 840
DB 781 YIKRVVPLATYSYGLMSFSLILLALPQOAYAVADASHGQIGALLVMTLFTLPFGYK 840
QY 840 YIKRVVPLATYSYGLMSFSLILLALPQOAYAVADASHGQIGALLVMTLFTLPFGYK 840
DB 840 YIKRVVPLATYSYGLMSFSLILLALPQOAYAVADASHGQIGALLVMTLFTLPFGYK 840

QY 841 TLISFPLMWLCYLLTLGEMVQEMAPMQVGRGDIIMAVAI FYPGVFDTIKMLLAVL 900
DB 841 TLISFPLMWLCYLLTLGEMVQEMAPMQVGRGDIIMAVAI FYPGVFDTIKMLLAVL 900
QY 901 GPAYLLKGLATVPYFVRAHALLMCTMARHLAGRYVQMLALLAGRTGTITYHILPM 960
DB 901 GPAYLLKGLATVPYFVRAHALLMCTMARHLAGRYVQMLALLAGRTGTITYHILPM 960
QY 961 SDMAASGLDLAVAEPIIFSPMEKKVIVWGAETAACDILHGLPVSARLGRVLLGPAD 1020
DB 961 SDMAASGLDLAVAEPIIFSPMEKKVIVWGAETAACDILHGLPVSARLGRVLLGPAD 1020
QY 1021 GYTSKMSLLABITAYAOQTRGLGTVIVSMTRDXTQAGEIQVLSVTYOSFLGTSISG 1080
DB 1021 GYTSKMSLLABITAYAOQTRGLGTVIVSMTRDXTQAGEIQVLSVTYOSFLGTSISG 1080
QY 1077 VCMYVHAGSKTTLAAPKGPITOMYTNVQDLYGMPKBPGRASLTPCTGSSDLVLYTRH 1136
DB 1077 VCMYVHAGSKTTLAAPKGPITOMYTNVQDLYGMPKBPGRASLTPCTGSSDLVLYTRH 1136
QY 1141 ADVIPARRRGDKGALLSPRLTLKSSGGPVLCPRGHAGVFPRAAVCSRGVAKSIDPI 1200
DB 1141 ADVIPARRRGDKGALLSPRLTLKSSGGPVLCPRGHAGVFPRAAVCSRGVAKSIDPI 1200
QY 1201 PVETLIDYTRSPFSDNSTPPAVPQTYQGYLHABTSGSKSTKVAVAAQGYKVLVNP 1260
DB 1201 PVETLIDYTRSPFSDNSTPPAVPQTYQGYLHABTSGSKSTKVAVAAQGYKVLVNP 1260
QY 1261 SVAAATLGFAYVSKAHGIDPNIRTVGRTITGAPATYSYSGKFLADGGSGGAYIICD 1316
DB 1261 SVAAATLGFAYVSKAHGIDPNIRTVGRTITGAPATYSYSGKFLADGGSGGAYIICD 1316
QY 1321 ECHAVDSTTLIGIVLDOAETAGVRLTVLATATPGSGTTPHPIIEVALGOEIEIPFY 1380
DB 1321 ECHAVDSTTLIGIVLDOAETAGVRLTVLATATPGSGTTPHPIIEVALGOEIEIPFY 1380
QY 1377 GRAIPLEAIRGRHLIFCHSKKKCDELAALRGMLNSVAYYRGLDVSYIPTQSDVVVA 1440
DB 1377 GRAIPLEAIRGRHLIFCHSKKKCDELAALRGMLNSVAYYRGLDVSYIPTQSDVVVA 1440
QY 1441 TDALMTGYGDPDSYIDCNVAATQVVDPSLDPTFTITTOIVQDVAVSORGRGRRL 1500
DB 1441 TDALMTGYGDPDSYIDCNVAATQVVDPSLDPTFTITTOIVQDVAVSORGRGRRL 1500
QY 1501 GIYRVYSTERASGMFDSVYLCECYDAGAAWELTPESETTVLRAVFPPTGYPVODHLE 1560
DB 1501 GIYRVYSTERASGMFDSVYLCECYDAGAAWELTPESETTVLRAVFPPTGYPVODHLE 1560
QY 1557 FMSVSTGLTHIDAFHLSQTKQAGDNFYLVAQATVCARAQAPPSMDQMCKILRKP 1616
DB 1557 FMSVSTGLTHIDAFHLSQTKQAGDNFYLVAQATVCARAQAPPSMDQMCKILRKP 1616
QY 1621 TLVGPPLLYRIGSYTNEVTLTHPVTKIATCOMADLEMTSTWVLAGVLAVAAYCILA 1680
DB 1621 TLVGPPLLYRIGSYTNEVTLTHPVTKIATCOMADLEMTSTWVLAGVLAVAAYCILA 1680
QY 1617 TLHGPTPLLYRIGSYTNEVTLTHPVTKIATCOMADLEMTSTWVLAGVLAVAAYCILA 1676
DB 1617 TLHGPTPLLYRIGSYTNEVTLTHPVTKIATCOMADLEMTSTWVLAGVLAVAAYCILA 1676
QY 1681 TGCVCIGRLHINOAVVAPDKEVLYEAFDEMEECASRAALTEEGORIEMLSKIQGL 1740
DB 1681 TGCVCIGRLHINOAVVAPDKEVLYEAFDEMEECASRAALTEEGORIEMLSKIQGL 1740
QY 1741 QOASQOADIOPTYOASWPKEOPFAKHMNFIISGIOYLALGISTLPGNPAVAMMAFSA 1800
DB 1741 QOASQOADIOPTYOASWPKEOPFAKHMNFIISGIOYLALGISTLPGNPAVAMMAFSA 1800
QY 1737 QATKQAEAAAVVSVSKWALLETFAKHMNFIISGIOYLALGISTLPGNPAVAMMAFSA 1796
DB 1737 QATKQAEAAAVVSVSKWALLETFAKHMNFIISGIOYLALGISTLPGNPAVAMMAFSA 1796
QY 1801 LNSPSTSTITLILNIGWLASQIAPRAGATGFVSGVGAUVGSIIGLKVVDLILAGG 1860
DB 1801 LNSPSTSTITLILNIGWLASQIAPRAGATGFVSGVGAUVGSIIGLKVVDLILAGG 1860
QY 1797 ITSPLTOSTLLENILGCVAAQOLAPPSAASFAVAGAAGAVGSIIGLKVVDLILAGG 1856
DB 1797 ITSPLTOSTLLENILGCVAAQOLAPPSAASFAVAGAAGAVGSIIGLKVVDLILAGG 1856
QY 1861 AGISGALVAFKIMSSEKPSMEDVNLPGIILPPGALVGVICGAILLRHVSGEGEAVQM 1920
DB 1861 AGISGALVAFKIMSSEKPSMEDVNLPGIILPPGALVGVICGAILLRHVSGEGEAVQM 1920
QY 1857 AGVAGLAVFKIMSSEKPSMEDVNLPGIILPPGALVGVICGAILLRHVSGEGEAVQM 1916
DB 1857 AGVAGLAVFKIMSSEKPSMEDVNLPGIILPPGALVGVICGAILLRHVSGEGEAVQM 1916

1921 NBLIAFASRGHVAPTHVYTESDASQRTQLLSLTITSLRLRHNMITECP1PCGSGW 1980
1917 NBLIAFASRGHVAPTHVYTESDASQRTQLLSLTITSLRLRHNMITECP1PCGSGW 1976
1981 LADVDMVCTITLTDFFKNMLTSKLPKMPGLPVVSCGKYGVAGTGIMTRCPGANIS 2040
1977 LADVDMVCTITLTDFFKNMLTSKLPKMPGLPVVSCGKYGVAGTGIMTRCPGANIS 2036
2041 GNVRLGSMRITGPTGMINIWOGEFPIINCTEGOCVKKPAPNPFVALIWRVAASFEYAVTQH 2100
2037 GHVKNCSMRIVGKTSNTWHGTFPIINATYTGCTSPANVSRALMRVAEEYVEVTR 2096
2101 GSYHYITGTTDLKVPYCOLPSEPFSSWDGVOIHFAPTPKPFDEVSFCGLNSFV 2160
2097 GPHVYTGTTDLKVPYCOLPSEPFSSWDGVOIHFAPTPKPFDEVSFCGLNSFV 2156
2161 GSOLPCEPDVPLVMSMLTDPSTHIAETAKRLIARSGPSEASASASQLSAPSLATCT 2220
2157 GSOLPCEPDVPLVMSMLTDPSTHIAETAKRLIARSGPSEASASASQLSAPSLATCT 2216
2221 THGKAVDVMVADNLF---MGSDVTRISGSKVVLVLSIDPWEERSDLEPSIPEYML 2276
2217 THHVSADALIEANLIMROEMGNITRVSEENKVVLDSFDPLRAEDEREVSVPAILR 2276
2277 PKKFPFPAIPAMAPDYNPPLVSWIKRPIYQAPATVACALPPRKTPTPPRRRRRTVGLS 2336
2277 KKKFPFPAIPAMAPDYNPPLVSWIKRPIYQAPATVACALPPRKTPTPPRRRRRTVGLS 2336
2337 EDSIGDALQOLAKSFGQPPSGDSLSTGAGADSQSOTP-PDELA-----LSETGIS 2390
2337 ESSVSSALAEIATKTG-----SSSSAVDSGTATALPDQASDGDGKSDIVESYS 2386
2391 SMPLEGEIGDDPLBEQVBPQPPGCVAAPESDSGSSTCEE-DBSVCCSMYSMT 2449
2387 SMPLEGEIGDDPLBEQVBPQPPGCVAAPESDSGSSTCEE-DBSVCCSMYSMT 2426
2450 GALITPCSPREEKLPINPLSNLSLRHNVKCTTTSASLRKAKVFPDMQVLDYSYDV 2509
2427 GALITPCSPREEKLPINPLSNLSLRHNVKCTTTSASLRKAKVFPDMQVLDYSYDV 2486
2510 LKDIKLAASKVTAARLITMEACQLTPHSAKSKYGFAGKEVSLSGRAVNHISKVMDL 2569
2487 LKEMKAKASTVAKKLITVEACKLTPHSAKSKYGFAGKEVSLSGRAVNHISKVMDL 2546
2570 EDSERPITPTIAKNEVFCVDPKGGKKAARLIVYDVGRCVCKALVDITQKLPQAVM 2629
2547 EDITVPIIDITIAKNEVFCVDPKGGKKAARLIVYDVGRCVCKALVDITQKLPQAVM 2606
2630 GASVGFQSPAPQVEFLKAMAEKDPMGFSYDTRCFDSTVTERDITREESTIRACSLPE 2689
2607 GSSVGFQSPAPQVEFLKAMAEKDPMGFSYDTRCFDSTVTERDITREESTIRACSLPE 2666
2690 EAHYAIHSLTERLYVGGFPMFNSKQTCGRRCASGVLTSMGNTITCYVKALAAKAA 2749
2667 EAHYAIHSLTERLYVGGFPMFNSKQTCGRRCASGVLTSMGNTITCYVKALAAKAA 2726
2750 IIAPTMLVCGDVLVVSQGTEDERNLRAFTAEATRYSAAPGDPREYDELITSS 2809
2727 LDDCTMLVCGDVLVVSQGTEDERNLRAFTAEATRYSAAPGDPREYDELITSS 2786
2810 SNVSVAGPQGRRRYVLTDPPTPIARAMEVTRHSPVNSWGLNIQVAPTIAMAVLMT 2869
2787 SNVSVAGPQGRRRYVLTDPPTPIARAMEVTRHSPVNSWGLNIQVAPTIAMAVLMT 2846
2847 HFFSILIAOELEKALDCOYIGACYSIEPLDLPQIIEHLHGSASFSLHYSPEINRVAS 2906
2870 HFFSILIAOELEKALDCOYIGACYSIEPLDLPQIIEHLHGSASFSLHYSPEINRVAS 2929
2930 ALRKXGAPLRAKMAKRAVRAVSLISRGRAVCGVLYNMVAKTKLTPLEAALLD 2989
2907 CRRKXGAPLRAKMAKRAVRAVSLISRGRAVCGVLYNMVAKTKLTPLEAALLD 2966
2990 SSMFTVAGAGGDIYHSVSRARPLLFGLLLFVGVGLFLPAR 3033

Db 2967 SCMFVAGYGGDIYHSVSRARPLLFGLLLFVGVGLFLPAR 3010

RESULT 8
US-08-324-977-14
; Sequence 14, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKE, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3010 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-324-977-14

Query Match 74.3%; Score 11984.5; DB 1; Length 3010;
Best Local Similarity 71.5%; Pred No. 0;
Matches 2175; Conservative 352; Mismatches 472; Indels 45; Gaps 8;
1 MSTNPKPORKTKRNTNRBPQDVKFPGGQIVGVYLLPRGRGLGVRAKTKTSRSQPRG 60

Db 1 MSTNPKQRKTKRNTNRNRPDVKFPGGGQIVGGVYLLPRRGPRILGVAPRKRTERSERQPRG 60
Qy 61 RRODIPKDRBSTGSKMGKPGYPMPLYGNEGLGNAGMLLSRGRSPSMGPRDPHRRSRNMG 120
Db 61 RRODIPKARPEGRITMAQPGYPMPLYGNEGLGNAGMLLSRGRSPSMGPRDPHRRSRNMG 120
Qy 121 KVIDTLTCGFADLMGYIPVVGAPLGVARALAHGVRVLEDDGVNPAATNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPVVGAPLGVARALAHGVRVLEDDGVNPAATNLPGCSFSIFLLA 180
Qy 181 LLSGCTTPVSAABEYKXNISTGYMTNDCNTDSITMOLAALVHAPGVCPCKEKVNAQCWT 240
Db 181 LLSGCTTPVSAABEYKXNISTGYMTNDCNTDSITMOLAALVHAPGVCPCKEKVNAQCWT 240
Qy 241 PVSPNVAVQPRGALTOGLRTHIDMVMSATLCSALVYDLCGGVMLAQMFIYSPQHMF 300
Db 241 ALVPTLLAARNVTIPTTIRRHVDLLVGAALFCSAMVYGDLCGSFVLVSOLFETSPRRHVT 300
Qy 301 VQDCNCSYIGGTTGHRMADMNMNNSPTATMLIAYMRVEVYIDIIISGHWGMFGLA 360
Db 301 LQDCNCSYIGGTHVSGHRMADMNMNNSPTALVVSQILRIPOAVVDVMAAHGVLAGLA 360
Qy 361 YFSGOGAMAVVYVLLLAAGVDAETHVGGSAOQTGRLTSLPDMGRKOTLOLVNNGSW 420
Db 361 YFSGOGAMAVVYVLLLAAGVDAETHVGGSAOQTGRLTSLPDMGRKOTLOLVNNGSW 420
Qy 421 HINRTALNCNDSLHTGFIAFLFYTHSFNSGCEPRMSACRSIEAFVGMGALOYEDNVTN 480
Db 421 HINRTALNCNDSLHTGFIAFLFYTHSFNSGCEPRMSACRSIEAFVGMGALOYEDNVTN 480
Qy 481 PEDMRPYCMHYPPRQCGVSAKTVCGPVYCFTPSPVVVGTDRLAGATYTMGENEDTVL 540
Db 481 PEDMRPYCMHYPPRQCGVSAKTVCGPVYCFTPSPVVVGTDRLAGATYTMGENEDTVL 540
Qy 541 LNSRPREPLSGMFCGTMMNNSGYTTTCGAPRCRTRADNASTDLCPTDCRKRKPDITLYK 600
Db 541 LNSRPREPLSGMFCGTMMNNSGYTTTCGAPRCRTRADNASTDLCPTDCRKRKPDITLYK 600
Qy 597 CGSGPMLTPRCMDVYFRLMHYFCTVNTFTFKRMVYGVSEHRLNAACNMTRECRDLED 656
Db 597 CGSGPMLTPRCMDVYFRLMHYFCTVNTFTFKRMVYGVSEHRLNAACNMTRECRDLED 656
Qy 601 CGSGPMLTPRCMDVYFRLMHYFCTVNTFTFKRMVYGVSEHRLNAACNMTRECRDLED 656
Db 601 CGSGPMLTPRCMDVYFRLMHYFCTVNTFTFKRMVYGVSEHRLNAACNMTRECRDLED 656
Qy 661 RDRSOLSPLLHSTTEWAILPCSYSDLPALSTGLHLHQNIVDVQFMYGLSPALTKYIVRW 720
Db 661 RDRSOLSPLLHSTTEWAILPCSYSDLPALSTGLHLHQNIVDVQFMYGLSPALTKYIVRW 720
Qy 721 EWTLLFLLADARVACIMMILLGQAEALAEKLVLIHAASAASCMGFYFIYFPAAM 780
Db 721 EWTLLFLLADARVACIMMILLGQAEALAEKLVLIHAASAASCMGFYFIYFPAAM 780
Qy 777 YIKGRVPLATYSLTGLMSFSLILLALPOOAYAYDASVHQIGNALMLITLFTLPFGYK 840
Db 777 YIKGRVPLATYSLTGLMSFSLILLALPOOAYAYDASVHQIGNALMLITLFTLPFGYK 840
Qy 841 TLLSRFLMWTLYLITLGEAMVQEMAPMVOVGRGDIIMAVALFYPGVVFDTKMLIAVL 900
Db 841 TLLSRFLMWTLYLITLGEAMVQEMAPMVOVGRGDIIMAVALFYPGVVFDTKMLIAVL 900
Qy 901 GPVALLGALTRVYFTRAHALLMCMABHLAGRIVOMALLALGNTCTYTYDHLTPM 960
Db 901 GPVALLGALTRVYFTRAHALLMCMABHLAGRIVOMALLALGNTCTYTYDHLTPM 960
Qy 961 SDMAASGLRDIAVNEPIIFSPMEKVIYVGAETPAACGDIILHGLPVARLARREVLLGPAD 1020
Db 961 SDMAASGLRDIAVNEPIIFSPMEKVIYVGAETPAACGDIILHGLPVARLARREVLLGPAD 1020
Qy 1021 GYTSKMSLLAPITAYVQOTRGLLIGTIVSMGTGDKTEQAGEIQVISTVTQSPFGTSISG 1080
Db 1021 GYTSKMSLLAPITAYVQOTRGLLIGTIVSMGTGDKTEQAGEIQVISTVTQSPFGTSISG 1080
Qy 1081 VLMVVYHAGNKTLASGRPYTOMYSABGLVGMWBPBPGTKSLSEPTCCGAVDLVYTRN 1140
Db 1081 VLMVVYHAGNKTLASGRPYTOMYSABGLVGMWBPBPGTKSLSEPTCCGAVDLVYTRN 1140
Qy 1077 VCMVTVHAGSKTLAARGPITOMYTNVDDLVGMWPKRPGARSLTPCTCGSSDLVYTRH 1136
Db 1077 VCMVTVHAGSKTLAARGPITOMYTNVDDLVGMWPKRPGARSLTPCTCGSSDLVYTRH 1136

Qy 1141 ADVIPARRRGDKRGALLSPRLSTLKGSSGGVLCPRGHAVGVFPAAVCSRGVAKSIDFI 1200
Db 1137 ADVIPARRRGDKRGALLSPRLSTLKGSSGGVLCPRGHAVGVFPAAVCSRGVAKSIDFI 1196
Qy 1201 PVETLIDVTRSPFSDNSTPBPVQOTYGVYHAAPTSKSKTKVPAVYAAQGYKVLVLP 1260
Db 1197 PVESMETTKRSVPFIDNSPPPAVPOSQVVAHHAAPTSKSKTKVPAVYAAQGYKVLVLP 1256
Qy 1261 SVAAATLGRAYLSKXHGIPNPIRTGVRTTGAPEITTYSTYGFELADGGCAGAYIIICD 1320
Db 1257 SVAAATLGRAYLSKXHGIPNPIRTGVRTTGAPEITTYSTYGFELADGGCAGAYIIICD 1316
Qy 1321 ECHAAVDSITLIGITVLDOAETAGVRLTVLATAPPGSVTTPHPNIEEVALQOGEIIPFY 1380
Db 1317 ECHAAVDSITLIGITVLDOAETAGVRLTVLATAPPGSVTTPHPNIEEVALQOGEIIPFY 1376
Qy 1381 GRAIPPLSYKGRHLIFCHSKKKCBELAAALRGMLNSVAYYRGIDVYIPIQGVVYVA 1440
Db 1377 GRAIPPLSYKGRHLIFCHSKKKCBELAAALRGMLNSVAYYRGIDVYIPIQGVVYVA 1436
Qy 1441 TDALMTGYTGFDSYIDCNVAATOVVDSLDPTFTTQIVPQDAVSRSGRGTGRGL 1500
Db 1437 TDALMTGYTGFDSYIDCNVAATOVVDSLDPTFTTQIVPQDAVSRSGRGTGRGL 1496
Qy 1501 GIYRYVSTGERASGMFDSVYLCECYDAGAAWELTPESETTVLRAVYFNTPGLPVCODHLE 1560
Db 1497 GIYRYVSTGERASGMFDSVYLCECYDAGAAWELTPESETTVLRAVYFNTPGLPVCODHLE 1556
Qy 1561 FWEAVFTGLTHIDAFHSQTKOSGENFAYLIAVQATVCARAKAPPSMDVMKCLTRLP 1620
Db 1557 FWEAVFTGLTHIDAFHSQTKOSGENFAYLIAVQATVCARAKAPPSMDVMKCLTRLP 1616
Qy 1621 TLVGPPLLYRGLSGVSTNEVTLNHPVTKYATCOMQDLEVMSTTWLAGVLAAYACLA 1680
Db 1617 TLVGPPLLYRGLSGVSTNEVTLNHPVTKYATCOMQDLEVMSTTWLAGVLAAYACLA 1676
Qy 1681 TGCVCIIIGRLHINQAVVAPADKEVLYEAFDEMEBCASRAALIEBQRIAEMLSKYIOGLL 1740
Db 1677 TGCVCIIIGRLHINQAVVAPADKEVLYEAFDEMEBCASRAALIEBQRIAEMLSKYIOGLL 1736
Qy 1741 QOASKOADIPTVOASMPKVEQFNAKHMNFIISGITYLAGISTIPGNPVAVSMAAFSA 1800
Db 1737 QOASKOADIPTVOASMPKVEQFNAKHMNFIISGITYLAGISTIPGNPVAVSMAAFSA 1796
Qy 1801 LTPSLSTFTIILNLLGWLASQIAPPGATGFVVSGLVGAIVGSIIGLKVVDLIAGY 1860
Db 1797 LTPSLSTFTIILNLLGWLASQIAPPGATGFVVSGLVGAIVGSIIGLKVVDLIAGY 1856
Qy 1861 AGISGALVAFKIMSGEKSPMEDVNLPGILSPGALVVGVIICAILLRHVGBGAVQWM 1920
Db 1857 AGISGALVAFKIMSGEKSPMEDVNLPGILSPGALVVGVIICAILLRHVGBGAVQWM 1916
Qy 1921 NRIAPASRGHVAAPHYTTESDASQVYQULIGSITISLRLRNMTTEDPIPCGSGN 1980
Db 1917 NRIAPASRGHVAAPHYTTESDASQVYQULIGSITISLRLRNMTTEDPIPCGSGN 1976
Qy 1981 LRDVMDWCTIITLDFNMULTSLPPMPGLPVSGOKYGVWAGGIGITTRCPCGANIS 2040
Db 1977 LRDVMDWCTIITLDFNMULTSLPPMPGLPVSGOKYGVWAGGIGITTRCPCGANIS 2036
Qy 2041 GNVRLSGMRITGPKTCMNIWQTFPINCYTEGQVCPKAPNPKVAIWRVAASEYAEVTOH 2100
Db 2037 GNVRLSGMRITGPKTCMNIWQTFPINCYTEGQVCPKAPNPKVAIWRVAASEYAEVTOH 2096
Qy 2101 GSYHYITGLITLNNLKVPQOLBEPFESVWDVQVQIHRFAPTPKPFREDEVSFCGLNSFVY 2160
Db 2097 GSYHYITGLITLNNLKVPQOLBEPFESVWDVQVQIHRFAPTPKPFREDEVSFCGLNSFVY 2156
Qy 2161 GSOLPCDPEPPTDVYMSMLTDPSSHITAEFARLARSGSPSSASSASQLSAPSLRATCT 2220
Db 2157 GSOLPCDPEPPTDVYMSMLTDPSSHITAEFARLARSGSPSSASSASQLSAPSLRATCT 2216

QY	2221	THGKAYVDVMDANF-----NGDVTRESSSKVVLVDSLDPMBEBSDBLEPSIPESEML	2276
Db	2217	THHVSDDDLLEANLIMRWQEMGNGINTRVESENKRVVLDSFDPLEAEBDEBEVSYPAEILR	2276
QY	2277	PKKFFPPALPMABPDVNPPLVESWIKRDYOPATVAGCALPPPKRTPTPPRRRRYVGLS	2336
Db	2277	KSKEFPAAWPIWAPEDVNPPLLESKMDDYVPVHGCPLPLPITAPPIPPRRRRYVLT	2336
QY	2337	EDSIGDALQOLAISFGQPPPSGDSGLSTGAGADSQSTP-PEELA-----LSEFGSIS	2390
Db	2337	ESSVSASLAELATKTFG-----SSESSAADSCTATLALPOASDDGDKGSDVESYS	2386
QY	2391	SMPLEGLGDPDLEPEQVEPQPPQGGVAAPGSSSGSWKSCSE-DDSVCCSMYSWT	2449
Db	2387	SMPLEGGPDPDL-----SDGWSWTSVEBASDVVCCSMYSWT	2426
QY	2450	GALITPCSPREEKLPINPLNSILRHYNHKVCTTTKASLPAKKYTFPDMOVLDYSYDSV	2509
Db	2427	GALLITPCAESKLPINLNSILRHHNMVYATTSRSAGLRQKQYTFPRLQVLDHTRDV	2486
QY	2510	LKDIKLAASKYTARLLTMEBAQCLTPPHSASKYFGKAEVRSLSGRAVHIKSWKDL	2569
Db	2487	LKEMKAKASTYAKULSYEBAQCLTPPHSASKFGYGAQDVRLNSKRVNHHSVWDL	2546
QY	2570	EDSEFPITTTIMAKNEVCPDPTGKGKAAALYYPDLCVRCERKALYDTOKLPOAVM	2629
Db	2547	EDYVPIPTTTIMAKNEVCPQPEKGGRKPARLIVFPDLGVAVCEKMLYDVVSTLPOVM	2606
QY	2630	GASVGFQYSPARVBEFLIKAAEKKDPMGFSDYDRCPSTYTEEDITEESIYACSLPE	2689
Db	2607	GSSTGFQYSPQARVEFLVNTKSKKNPMGFSYDRCPSTYTENDIREESIYQCCDAP	2666
QY	2690	EAHTAISHLERLVYVGGPMFNSKQOTCYRRASGAVLTSMGNTITCYVVALAACRAAG	2749
Db	2667	EAQQAISLTERLIVIGBLTMSKQNCYRCRASGVLTSCGNTITCYLKASAAACRAK	2726
QY	2750	IIAPTMLVCGDVLVVISSEOSTEEDERNLRAFTEAMTRYSAPEGDPPEPEYDLITSCS	2809
Db	2727	LQDCMTLVNGDVLVVICESAGTQEDBASLRVFTAMTRYSAPEGDPPEYDELITSCS	2786
QY	2810	SNVSVALGPOGRRRYLTRDPTTPIARAMETVSHSPNSWGLNITQAPITMARMLMT	2869
Db	2787	SNVSVAHDASGRVYLLTRDPTTPIARAMETAHTTPNSWGLNITVAPITMARMLMT	2846
QY	2870	HFPSILMAQDTLDONLFNEMYGAVYSVPLDPAIIERLHLDLAFSLHTTYPHBLTVAS	2929
Db	2847	HFPSILMAQDDEKRALDCQIYGACYSIEPLDPOIIRLHLSAFSLHSYPSGIRNVAS	2906
QY	2930	ALRLKIGAPPLPAWKSARAVASLISRCGRDAVCGRYLFNNAVTKLKLTPLEPARLLD	2989
Db	2907	CLRLKIGVPLPLVWRHARSVARLLSQCGRAATCGKYLFNNAVTKLKLTPIPASRLD	2966
QY	2990	SSWTVVAGAGGDITHSVGRAPRLILFGLLLFPVGVGLFLLIPAR	3033
Db	2967	SGWFAVAGSGDIYHSLSRARPRFMCLLLLSGVGIYLLPNR	3010
RESULT 9			
US-08-384-616-2			
; Sequence 2, Application US/08384616			
; Patent No. 5847101			
; GENERAL INFORMATION:			
; APPLICANT: OKAYAMA, Hiroto			
; APPLICANT: FUKU, Ieao			
; APPLICANT: MORI, Chisato			
; APPLICANT: TAKAMIZAWA, Akehisa			
; TITLE OF INVENTION: YOSHIDA, Iwao			
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC			
; NUMBER OF SEQUENCES: CDNA AND ANTIGEN POLYPEPTIDE			
; CORRESPONDENCE ADDRESS: 50			
; ADDRESS: Armstrong, Westernman, Hattori, Mclelland &			
; ADDRESS: Naughton			

```

1 STREET: 1725 K St. N.W. Suite 1000
2 CITY: Washington
3 STATE: D.C.
4 COUNTRY: U.S.A.
5 ZIP: 20006
6 COMPUTER READABLE FORM:
7 MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
8 COMPUTER: IBM PC compatible
9 OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
10 SOFTWARE: ASCII
11 CURRENT APPLICATION DATA:
12 APPLICATION NUMBER: US/08/384,616
13 FILING DATE:
14 CLASSIFICATION: 424
15 PRIOR APPLICATION DATA:
16 APPLICATION NUMBER: US 07/769,996
17 FILING DATE: 02-OCT-1991
18 APPLICATION NUMBER: JP 2-167466
19 FILING DATE: 25-JUN-1990
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER: JP 2-30921
22 FILING DATE: 31-AUG-1990
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: JP 2-305605
25 FILING DATE: 09-NOV-1990
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER: US 07/635,451
28 FILING DATE: 28-DEC-1990
29 ATTORNEY/AGENT INFORMATION:
30 NAME: Stevens-Smith, Theresa M.
31 REGISTRATION NUMBER: 36,281
32 REFERENCE/DOCKET NUMBER: 900703B
33 TELECOMMUNICATION INFORMATION:
34 TELEPHONE: (202) 659-2930
35 TELEFAX: (202) 887-0357
36 TEXEL: 440142
37 INFORMATION FOR SEQ ID NO: 2:
38 SEQUENCE CHARACTERISTICS:
39 LENGTH: 3010 amino acids
40 TYPE: amino acid
41 TOPOLOGY: linear
42 MOLECULE TYPE: protein
43 US-08-384-616-2
44
45 Query Match 74.3%; Score 11984.5; DB 2; Length 3010;
46 Best Local Similarity 71.5%; Pred. No. 0;
47 Matches 2175; Conservative 352; Mismatches 472; Indels 45; Gaps 8;
48
49 QY 1 MSTNPKPQRTKRNKTNRPQDVKFPGGQIVGVYLLPRRGPRLGVATRTKTSERSOPRG 60
50 DB 1 MSTNPKPQRTKRNKTNRPQDVKFPGGQIVGVYLLPRRGPRLGVATRTKTSERSOPRG 60
51 QY 61 RRQGIPIPKDRSTGKSMCKRGVPMPLYNBGLGMAWLLSRSGRPSMGPNDPHRSNNVG 120
52 DB 61 RRQGIPIPKDRSTGKSMCKRGVPMPLYNBGLGMAWLLSRSGRPSMGPNDPHRSNNVG 120
53 QY 121 KVIDTLTCGFDLMGVIPIVGAFLGAVARALAHGVRILEGVAIYATGNLGCSSFIFLLA 180
54 DB 121 KVIDTLTCGFDLMGVIPIVGAFLGAVARALAHGVRILEGVAIYATGNLGCSSFIFLLA 180
55 QY 181 ILSCTTPVSAAEVKNISTGYMTNDCTNDSITWOLQAALVHPGCVPCBEKVNASQCMW 240
56 DB 181 ILSCTTPVSAAEVKNISTGYMTNDCTNDSITWOLQAALVHPGCVPCBEKVNASQCMW 240
57 QY 241 PVSENVAVQRPALTOGLRTHIDMTVMSATLCSALTYGDI CGVMTLAAQMFIVSPQHMF 300
58 DB 241 ALPTTLAARVVIPTTIRRHVDLLVGAALFCSAMYVDLCGSVFLVSQLFRTSPRRHT 300
59 QY 301 VQDNCGISYGTITGHRMADMMNMNSPTATMTLAVAMRPEVYIIIDISAHGVMFGLA 360
60 DB 301 VQDNCGISYGTITGHRMADMMNMNSPTATMTLAVAMRPEVYIIIDISAHGVMFGLA 360
61 QY 361 VFSNKGMAVKAUVYILLAAQVDAIRTHVGGSSAAQTTGRILSLPDMGPBROKIOLVNTGSW 420
62 DB 361 VFSNKGMAVKAUVYILLAAQVDAIRTHVGGSSAAQTTGRILSLPDMGPBROKIOLVNTGSW 420

```

Db 361 YYSAGNMAKYLIVMLFAGVDGDTHTVGGAQAKTIRLIVSMFASGSPQIQLINTGSMW 420
Qy 421 HIRNTALNCDSLHTGFIASLFTTHSPNSSGCEPRMAGCSIEAFRGMALQYEDNVN 480
Db 421 HIRNTALNCDSLHTGFIASLFTTHSPNSSGCEPRMAGCSIEAFRGMALQYEDNVN -SS 478
Qy 481 PEDMRPYCMHYPPROCGVSAKTVCGPVYCTPSPVVVGTTDBRLGAPTYTGMENEDVFL 540
Db 479 RSDQRPVCMHYPPROCTIVPASEVCGPVYCTPSPVVVGTTDBRFVGTPTMKGNEEDVLL 538
Qy 541 LNSTRPPLSGWFGCTMNMSSGYTKTCGAPPCRTBADFNASTDLICPTDCRKPDDTYLK 600
Db 539 LNNTRPPOGNWFGCTMNMSSGYTKTCGAPPCRTBADFNASTDLICPTDCRKPDDTYLK 596
Qy 601 CGSGPMLTPRCILDYPRMLMHPCTVYNTI PKIRMYVGVGVEHRLTAACNTRDRCDLED 660
Db 597 CGSGPMLTPRCMDYPRMLMHPCTVYNTI PKIRMYVGVGVEHRLTAACNTRDRCDLED 656
Qy 661 RDRSOLSPLLHSTTEWAILPCSYSDDLPAIYSTGLHLHQNIVDVQFMYSPLATKYIVM 720
Db 657 RDRSOLSPLLHSTTEWAILPCSYSDDLPAIYSTGLHLHQNIVDVQFMYSPLATKYIVM 716
Qy 721 EWWILLFLIADARVCACTMMLILGQAEALAEKLVILHAASAACNGFLYFVIFVAAW 780
Db 717 EWWILLFLIADARVCACTMMLILGQAEALAEKLVILHAASAACNGFLYFVIFVAAW 776
Qy 781 YIKRVVPLATYSITGMSFSLIILALPOQAVYVADSVHGOIGALLVMTLETLTPGYK 840
Db 777 YIKRVVPLATYSITGMSFSLIILALPOQAVYVADSVHGOIGALLVMTLETLTPGYK 836
Qy 841 TLLSRFLMMLCYLLTLLEAMVOWEAPMVOYRGDGIIMAVAIYFPGVVEDITKMLAVL 900
Db 837 VFLARLIMWLYFTTRAEDADLHWIIPRLNARGGDAILLIMCAVHPELIDITKMLAIL 896
Qy 901 GPAYLLKALTRVYEFRAHALLMCTMARHLAAGRVQWALLALGMWGTYYIYDHLTPM 960
Db 897 GPAYLLKALTRVYEFRAHALLMCTMARHLAAGRVQWALLALGMWGTYYIYDHLTPM 956
Qy 961 SDMAASGLRLDAVNEPILFISPMKKYIWMGAETPAGDILHGLPVASARLGREVLGPAD 1020
Db 957 SDMAASGLRLDAVNEPILFISPMKKYIWMGAETPAGDILHGLPVASARLGREVLGPAD 1016
Qy 1021 GYTSKGMSLAPITAYAQOTRGLIGITIVSMGTGRDKEQAGEIQVLSVTYQSEFGLTISG 1080
Db 1017 GYTSKGMSLAPITAYAQOTRGLIGITIVSMGTGRDKEQAGEIQVLSVTYQSEFGLTISG 1076
Qy 1081 VLMVTYHGAQNKTLAAGRPVYQWYSSAEGDLVGMSPPGTKSLEPCTCGAVDLYLVTRN 1140
Db 1077 VLMVTYHGAQNKTLAAGRPVYQWYSSAEGDLVGMSPPGTKSLEPCTCGAVDLYLVTRN 1136
Qy 1141 ADVPARRDGKALISPRPLSTLKGSSGGPVLCPRGHAVGVPRRAVAVCGRGVAKSIDFI 1200
Db 1137 ADVPARRDGKALISPRPLSTLKGSSGGPVLCPRGHAVGVPRRAVAVCGRGVAKSIDFI 1196
Qy 1201 PVETLIDIVTRSPFSDNSTPBAVPOQYOVGYLHAAPTSGSKSTKVPAVYAAQGYKVLVLP 1260
Db 1197 PVETLIDIVTRSPFSDNSTPBAVPOQYOVGYLHAAPTSGSKSTKVPAVYAAQGYKVLVLP 1256
Qy 1261 SVAAITLFGAYMSKMGIDPNIKIRTVGVTITTTGAPVITSTYVKGFLADGCGSGGAYDITICD 1316
Db 1257 SVAAITLFGAYMSKMGIDPNIKIRTVGVTITTTGAPVITSTYVKGFLADGCGSGGAYDITICD 1316
Qy 1321 ECHAVDSTTILIGITVLDQAEATAGVRLTVLATATPPGSAVTTPHNTEVVALGQEGEIPFY 1380
Db 1317 ECHAVDSTTILIGITVLDQAEATAGVRLTVLATATPPGSAVTTPHNTEVVALGQEGEIPFY 1376
Qy 1381 GRAIPLSYIKGRHLIFCHSKKCDLAALRGWGLNSVAYYRGDLVSVIPTQGDVVVA 1440
Db 1377 GRAIPLSYIKGRHLIFCHSKKCDLAALRGWGLNSVAYYRGDLVSVIPTQGDVVVA 1436
Qy 1441 TDLMTGYTGDPSVIDCNAVTVQVNDPSLDPTTITTOIVPQDANVSRGRGTGRGL 1500
Db 1437 TDLMTGYTGDPSVIDCNAVTVQVNDPSLDPTTITTOIVPQDANVSRGRGTGRGL 1496
Qy 1501 GYRYVSTGERASGMFDSVVLCECYDAGAAWELPSETTVLRAFYPTPGLPYVCOHLE 1560
Db 1497 GYRYVSTGERASGMFDSVVLCECYDAGAAWELPSETTVLRAFYPTPGLPYVCOHLE 1556
Qy 1561 FWEAVFTGLTHIDAFHLSQTKOSGENFAVLTAYQATVCARAKAPPPSDVMWKLTRLKP 1620
Db 1557 FWEAVFTGLTHIDAFHLSQTKOSGENFAVLTAYQATVCARAKAPPPSDVMWKLTRLKP 1616
Qy 1621 TLVGPPLLYRLGASTNEVTLTHPTKYIATQMADLEVMSTWVLGAVLAANAACLA 1680
Db 1617 TLVGPPLLYRLGASTNEVTLTHPTKYIATQMADLEVMSTWVLGAVLAANAACLA 1676
Qy 1681 TGCVCIGRHLINOGAVVAPDKEVLYEAFDEKEECASRAALIEBGGRTAEMLSKXIQLL 1740
Db 1677 TGCVCIGRHLINOGAVVAPDKEVLYEAFDEKEECASRAALIEBGGRTAEMLSKXIQLL 1736
Qy 1741 QQASRQADIDPTVOASMPKVEQFWAKIMMNFISGIIQYLAGUSTLPGNPVAVASMAFSA 1800
Db 1737 QQASRQADIDPTVOASMPKVEQFWAKIMMNFISGIIQYLAGUSTLPGNPVAVASMAFSA 1796
Qy 1801 LNSPLSTTITLNLIGWHLAQIAPPAAGATGVVSGVGAAGVSGIGLKVLDILAGV 1860
Db 1797 LNSPLSTTITLNLIGWHLAQIAPPAAGATGVVSGVGAAGVSGIGLKVLDILAGV 1856
Qy 1861 AGISGALVAFKIMSEKSPMEDVNLPLGISPGALVGVICAAILRRHVSGEGAVQWM 1920
Db 1857 AGISGALVAFKIMSEKSPMEDVNLPLGISPGALVGVICAAILRRHVSGEGAVQWM 1916
Qy 1921 NRLIAPASRGNHAPTHVYTESDASQRYTOLLGSLITTSILRLHNMITEDECPICGSM 1980
Db 1917 NRLIAPASRGNHAPTHVYTESDASQRYTOLLGSLITTSILRLHNMITEDECPICGSM 1976
Qy 1981 LRDVMDWCTITLTDKRMNLTSLFPRMGLPVSQOKYKGMAGTGMITTRPCGANIS 2040
Db 1977 LRDVMDWCTITLTDKRMNLTSLFPRMGLPVSQOKYKGMAGTGMITTRPCGANIS 2036
Qy 2041 GNVRLGSMRITGPCKMNIWOSTFPINCYTEGQCPKPAKPKVAILMRYAASEVAVTQH 2100
Db 2037 GNVRLGSMRITGPCKMNIWOSTFPINCYTEGQCPKPAKPKVAILMRYAASEVAVTQH 2096
Qy 2101 GSRYHITGLITDNLKVPQCLPSPEFFSWDVQIHRFAFTPPKPFREDEVSFCVGLNSFVY 2160
Db 2097 GSRYHITGLITDNLKVPQCLPSPEFFSWDVQIHRFAFTPPKPFREDEVSFCVGLNSFVY 2156
Qy 2161 GSQCLDPEPDDVLYMSMLTDPSSHITAEFARLARLARGSPSSAASSASQSLAPSLACT 2220
Db 2157 GSQCLDPEPDDVLYMSMLTDPSSHITAEFARLARLARGSPSSAASSASQSLAPSLACT 2216
Qy 2221 THGKAVDVMVANI.F---MGGDVTRIESGSKVVVLDLDPMEVERSDLEPSISEVWL 2276
Db 2217 THGKAVDVMVANI.F---MGGDVTRIESGSKVVVLDLDPMEVERSDLEPSISEVWL 2276
Qy 2277 PKKREPPALPAMARPDYNPPLVESWKRBDYOPATYVAGCALPPPKTPTTPPPRRRTVGIS 2336
Db 2277 PKKREPPALPAMARPDYNPPLVESWKRBDYOPATYVAGCALPPPKTPTTPPPRRRTVGIS 2336
Qy 2337 EDSIGDALOQLIKSFQGPFPBGDSGLSTGAGAASSGPTP -PDELA-----LSEFGIS 2390
Db 2337 EDSIGDALOQLIKSFQGPFPBGDSGLSTGAGAASSGPTP -PDELA-----LSEFGIS 2386
Qy 2391 SMPPLEGELGDPDLEBQVEPQPPQGGVLAAGSDSGSWSCTSEE -DSDVSCSSMSYMT 2449
Db 2387 SMPPLEGELGDPDLEBQVEPQPPQGGVLAAGSDSGSWSCTSEE -DSDVSCSSMSYMT 2426
Qy 2450 GALITPCEPDEEKLPINPLNSILRYHNKRYCTTTSASLRAKKVTFDRMOVLDSYDSV 2509
Db 2427 GALITPCEPDEEKLPINPLNSILRYHNKRYCTTTSASLRAKKVTFDRMOVLDSYDSV 2486
Qy 2510 LKDIKLAASKVTAARLLIMEAQLTPPHSABSKYFGAKAEVNSLGRAVNHAKSVWKDL 2569
Db 2487 LKDIKLAASKVTAARLLIMEAQLTPPHSABSKYFGAKAEVNSLGRAVNHAKSVWKDL 2546

Qy	721	EWVLLPILLADAVCACTLMWLLILGQAEALTEKVLITHASASCSNGFLYVFVFPAAW	780
Db	717	EYVLLLEFLLADARVCACTMMWLLIAQEALELVLTNSASVGAHGILSFVFPAAW	776
Qy	781	YIKRVVPLATYSLTGLWSFSLLLLALPQCAVVAADASVHQCIGAALLVMTLFLYTGKX	840
Db	777	YIKRLVPGATYALYGVWPLLLLLALPRAVYMDREMAASCGGAVVGLVLLTSYKX	836
Qy	841	TLLSRFLMWLCYLLTLTGAVNOEMAPMQVNRDGIIMAVATFVPGVDPITKMLAVL	900
Db	837	VFLARLIMWLQYFTTRPADLHVMIPLNARGSDAILLLMCVAHPDLFDITKLLAIL	896
Qy	901	GPVLLKCALTRVVPFVAHAAHLLMCTMARILHAGRVYOMALLAGMTGYVTDHLLPM	960
Db	897	GPLMWLQGITRVPLFVAHQGLIHACMLVRVAGAHVQMAFMKLGALVTGYVYNHLLPL	956
Qy	961	SDMAASGLRDIAVAVEPIIFSPMEKXVIVMAEFTAAACGDIHGHVPASRLCREVLLPAD	1020
Db	957	RDWRPAGRDIAVAVEPVFSDEMKEKIIWKADTRACGDIILGLPVSARKKEILLPAD	1016
Qy	1021	GYSKGSWLLAPITAAVQOTRGLLGTIVSVMTGDKTEQAGEIYVLSTVTSFGLTSG	1080
Db	1017	SLEGRGLRLAPITAVSQOTRGLLGCITTSLTGRDKNQVEGEVQVSTATQSFATCVNG	1076
Qy	1081	VLMTVYHAGKXKTLAGSRGPYQWYSASBGLVWGPSPGKSLSEPTCCAVDILYVTRN	1140
Db	1077	VCWTVYHAGSKTLAAPGPITQWYVNDOLVWPKPGRASLTPCTCGSSDILYVTRH	1136
Qy	1141	ADVPARRGGRKGLASPRPLSTLKGSSGGPVLCPRHAAGVPRAAVCSRGVSKSIDFI	1200
Db	1137	ADVPVRRRGRSGSLSPREVSYLKSSGGPLCPFHAVGIRAAVCTRGVAKADVFL	1196
Qy	1201	PVELLDIVTRSPFSDNSTPRAPVQTYOVYLHAPTGSGSKTKVPVAAAGYKVLVLP	1260
Db	1197	PVESMETMRBPVFTDNSSPRAPVPSFQVAILHAPTGSGSKTKVPAAVAAGYKVLVLP	1256
Qy	1261	SVAATLGGAYLSKRAHGINIRGVRTVTGAPITTSYTKKFLADGGCAGADYIICD	1320
Db	1257	SVAATLGGAYWSKHAHDIPNIRGVRTITTGAPVATYSTYKFLADGGCSGAADIICD	1316
Qy	1321	ECHAVDSFTLIGIVLQDAETAGRLVLTATPFGSVTPPHNIEEVALGQGEIIPFY	1380
Db	1317	ECHSDSTTLIGIGIVLQDAETAGRLVLTATPPGSAVTPPHNIEEVALNSTGEIIPFY	1376
Qy	1381	GRAIPLSYIKGRHLIFCHSKKCCDELAAALRGGLNSVAAYYRGIDVSIPTQGDVVVVA	1440
Db	1377	GKAIPPIEAIROGRHLIFCHSKKCOELAAKLSGLGINAAVYRRGLDVSIVPTIDGVVVA	1436
Qy	1441	TDALMTGYTGPDSYIDCNNAVATQVDPSSLDPFTTITQIYVPDASVSQRRGTGRLL	1500
Db	1437	TDALMTGYTGPDSYIDCNCTQTVDPDSLDPFTFIETTVPODASVSQRRGRTGRRL	1496
Qy	1501	GIYRVYVSGBEASGPFDSVULCEQDACAAMAYELTPSTYRLAYFNTPELPLVCOHLE	1560
Db	1497	GIYRVYVGBEBSGPFDSVULCEQDACAAMAYELPATYSRLAYLNTPELPLVCOHLE	1556
Qy	1561	FMEAVFTGLTIDAHAFLSQTKOSGENFYLITAYQATVACARAKAPPPSDVMWKKLTLPK	1620
Db	1557	FMEASFGLTTHIDAFLSQTKQAGNPFYLYAYQATVACARAKAPPPSDVMWKKLIRKP	1616
Qy	1621	TLVGPBLLYRLGSVTNEVTLTHPVTKYIATCMQADLEVMTSTWVLAGVLAAVAACLA	1680
Db	1617	TLHGFTPLLYRLGAVQNEVTLTHPIITKYIMACMSADLEVMSTWVLVGVLAALAAACLT	1676
Qy	1681	TGCVCLITGRLLHINQAVVAPDKVLYEAFDEMEEBASAALLIEBQRIAEMLKSKIQGLL	1740
Db	1677	TGSVVIVIRIILSGRPAIVPRELLYQEFDEMEEBASHLPYIEGQMOLAEQFKKALGLL	1736
Qy	1741	QOASRKQADIDPTVQASWPKVQOPAKKMMNFIISGIQYLAGISTLPGNPAVASMMAFSA	1800
Db	1737	QYATYQAAAPAVVESKRBALETFPAKMMNFIISGIQYLAGISTLPGNPALASLMAFPAS	1796
Qy	1801	LTSPLSTSTILLNLIGWLASQIAPRAGATGFVWSGLGVAAGSIGKVLVDILAGYG	1860

Db	1797	ITSPLETTQSTLLEFNLISGWAAQLAPSPASAFAGAGIAGAAAGSISLQKVLVDILGYG	1856
Qy	1861	AGISGALVAEFKIMSGEKPEKSMEDVNNLLPGILSPGALVGVICAIIERRHYGPEGAVOM	1920
Db	1857	AGVAGALVAFKVMGEMPMSTEDLVNLLPALISPEALVGVVCAAILRRHGPGGAVOM	1916
Qy	1921	NRLIAFASRGNHVAPTHYVYESDASQRYTQLSGSLTISLRLRHMMITEDCP1PCGSG	1980
Db	1917	NRLIAFASRGNHVSPTHYVYESDAAAVTOILSSLTITOLLRHMQMINEDCSTPCSGS	1976
Qy	1981	LBDVMDVCTLLDFQKMWLSKLEPKKPGJPFVSCQKGYGVAMGATGIMTTRCCGANIS	2040
Db	1977	LBDVMDMCTVLTDFKMTWLOGSKLPLQJPGVPFSCQKGYGVAMGDMQTCECGAQT	2036
Qy	2041	GNVLGSMRITGPFKTCMMIWOETPINCYEGQCVCPAPNFKVAIWRVAASEAYEYOH	2100
Db	2037	GHVNGSGMRIIGPRTCSNTMHGTGPINAYTTGCTSPAPAPNYSALMRVAABEYVETRV	2096
Qy	2101	GSYHYITGLTNDLKVPCQLSPFEFSWVDVQIHRAPLPKPEFRDEVSFCVGLNSFV	2160
Db	2097	GDFHYVGMETDNDKCECQVPAPEPFEBVGVRIHRAAPACRPLIREEVTFQVGLNOYLV	2156
Qy	2161	GSOLPCPEPPTDVLNMDPSHITETARLARARSPSEASSSASOLSAPSLRATCT	2220
Db	2157	GSQLPCEPEPVVALTMLTDPHSITABETARRIARSGSPSLSSSSASQLSAPSLKATCT	2216
Qy	2221	THGAYVVDWMDANLTF---MGCVNTR1ESGSKVVLVDSLDPMWEERSDLEP1ISEYML	2276
Db	2217	THYVSPDADLLEANLWRQEMGNITVYESNKKVVLVDSFPDLAEDREHVSYPAILLR	2276
Qy	2277	PKKRPPEPALPMAPDPDYNPLVESWKRPDQOPATVACALPPEPKTTPPPRRRRYVGLS	2336
Db	2277	KSKEFPAMPPIWAPDPDYNPLLESWMDPDYVPVPHVCCPLPIKAPR1PEPPRRKRTVLT	2336
Qy	2337	EDS1GDMLQOLA1KSPQCPFPSSGSGSLTAGADSSQRP-PELA-----1SETGIS	2390
Db	2337	ESSVSSALAE1ATATTFG-----SSBSADVDSGTATALPPOASDDGDKGSDVESYS	2386
Qy	2391	SMPPLEGLGDPDLEPEQVEPQPPGCGVAAPGSDSGSWSTCSBE-DSVVCSCMSYMT	2449
Db	2387	SMPLEGERGDPD-----SDGSMVSYSEABEDVVCSCMSYMT	2426
Qy	2450	GALITPCSPBEKLPINPLNSNLLRYHNKYCTTTSKASLRKAKVYFDRMQVDSYDSV	2509
Db	2427	GALITPCABESKLPINALSNSLLRHHNMVYATTSRAGLRQKVTYFDRLOVDLDHYRDV	2486
Qy	2510	LKD1KLASKTYAALLMEBAQCLTTPHSARSKTGFKAKEFRSLSGAVNHIKSWMDL	2569
Db	2487	LKEMKAKASTYKAKLISVEBAKCTLTPHSARSKFGYAKDVRNLSRAVNHISHWMDL	2546
Qy	2570	EDSETP1PTTMAKNVEYCVPTGSGKKAARLIYVPLGVYVCEKMLYDTQCLPOAVM	2629
Db	2547	EDYTP1DITTIMAKNVEYCVQPEKGGKPARLIYFPLDGVYVCEKMLYDVSTLPOVVM	2606
Qy	2630	GASYGFOYSPQARVEFLKAWAEKKDPMGFSYDTRCEDSTVTEERDIRTESIYRACSLPE	2689
Db	2607	GSSYGFQYSPQARVEFLNMTKSKKNPMGFSYDTRCEDSTVTEERDIRVESIYQCCLAP	2666
Qy	2690	EATHAHSLTERLYVGGPMENSKOQTCYRRCRASGVLTSSMGNTITCYVVALAACAAAG	2749
Db	2667	EARQAKSLTERLYIGPLTNSKQNGCYRRCRASGVLTSSCGNTITCYLVAASACRAAK	2726
Qy	2750	I1APMLVACGDLVVISGQTEEDENLR1AFETAMRYSAPEGDP1PRPEYDELLITSCS	2809
Db	2727	LQDCMLNMGDLVVICESAGTQDPAASLR1FTAMR1YSAPEGDP1PRPEYDELLITSCS	2786
Qy	2810	SNVSVALGPOQRRRYLRLDPTTET1AABMETVSHSPVNSMGN1IOYAP1TAMRYMLMT	2869
Db	2787	SNVSVAHDASGRYYVYLTRDPTTTLARAME1AHTTPVNSMGN1IMVAP1TAMRYMLMT	2846
Qy	2870	HFPS1MAQDTLDONLNFEMTGAVYSPLDPA1IERLHGLDAPSLHTTYTPHELTRYAS	2929

Db 2847 HFFSILLAOEQLEKALDCQIYGACYSIEPLDLPQIIEIRLHGLSAFSLHSYSPGEINRVAS 2906
Qy 2930 ALRKIGAPRLRMKSRABVRASLIRGGRAAVCGYILFMVAKTKLKTPLREALLDL 2989
Db 2907 CRRKGVPLRWRHRAKSRVRLALSQGRALATCGYILFMVAKTKLKTPLPAASRLD 2966
Qy 2990 SSMFTYAGAGCDIYHSVSRAPRLILFGLLTFVGLFLLPAR 3033
Db 2967 SCMFVAGYSGCDIYHLSRARRPFWMLCLLLSVGIGYILLPR 3010

RESULT 11
US-08-904-686A-2
Sequence 2, Application US/08904686A
Patent No. 5998130
GENERAL INFORMATION:
APPLICANT: OKAYAMA, Hiroto
APPLICANT: FURE, Isao
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &
ADDRESS: Naughton
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/904,686A
FILING DATE: 01-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Mclelland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 900703G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3010 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLBULE TYPE: protein
US-08-904-686A-2
Query Match 74.3%; Score 11984.5; DB 2; Length 3010;
Best Local Similarity 71.5%; Pred. No. 0;
Matches 2175; Conservative 352; Mismatches 472; Indels 45; Gaps 8;

Qy 1 MSTNRPKPKTKRNNRNRRODYKPPGCGQIVGGVLLPFRGRGLGRATRKTSERSQPG 60
Db 1 MSTNRPKPKTKRNNRNRRODYKPPGCGQIVGGVLLPFRGRGLGRATRKTSERSQPG 60

Qy 61 RRQPIKORRSTGSKMGKPEYMPPLYNGELGMAGMLSPRSRSPMGPNDRHSRNVG 120
Db 61 RRQPIKORRSTGSKMGKPEYMPPLYNGELGMAGMLSPRSRSPMGPNDRHSRNVG 120

Qy 121 KYIDLTTCGFADLMGYIPVGAAPLGGVARALAHGVRLVEDGVNPFATGNI PGCSFSIFLLA 180
Db 121 KYIDLTTCGFADLMGYIPVGAAPLGGVARALAHGVRLVEDGVNPFATGNI PGCSFSIFLLA 180

Qy 181 LLSCTTPVSAAEVKNISITGVVNTDCTNDSITWQLQAALVHVPQVCPCEKVGNAQCMI 240
Db 181 LLSCTTPVSAAEVKNISITGVVNTDCTNDSITWQLQAALVHVPQVCPCEKVGNAQCMI 240

Qy 241 PVSFVAVQRPALITQGLRTHIDMYVMSATLCSALYVGDLCGVMLAQMFTVSPQHMF 300
Db 241 ALTPILANANVTIPIPTTIRRHVDLVGAAPFCSAMYVGLDGSVFLVSGLPFSPRRHYT 300

Qy 301 VDDCNCSITPGTITGHRNAMDMMNWSPTATMTILAVNRVPEVIIDISGAHGMVFGIA 360
Db 301 LQDCNCSITPGHVSCHRMAMDMMNWSPTATMTILAVNRVPEVIIDISGAHGMVFGIA 360

Qy 361 YFSMGAMAKVIVILLAAVDARTHTVCGSAAOQTGRLSLFDMGPRKIDLVNNGSM 420
Db 361 YFSMGAMAKVIVILLAAVDARTHTVCGSAAOQTGRLSLFDMGPRKIDLVNNGSM 420

Qy 421 HINRTALNCNDSIAHGFIALSFYTHSFNSSGCEBMSACRSIEARVFGALQYEDNVN 480
Db 421 HINRTALNCNDSIAHGFIALSFYTHSFNSSGCEBMSACRSIEARVFGALQYEDNVN 480

Qy 481 PEDMRPYCHNPPROCGVVSACTVCGPVYCFPPSPVNVGTTDRDLGAPYTWGBENTDVL 540
Db 479 RSDQRPYCHNPPROCTIVPASEVCGPVYCFPPSPVNVGTTDRDLGAPYTWGBENTDVL 538

Qy 541 LNSTRPPLGSMFGCTMMNSSGYTKTCGAPPCGTRADFNASTDLDLCPDCEFGHPDTTYAK 600
Db 539 LNSTRPPLGSMFGCTMMNSSGYTKTCGAPPCGTRADFNASTDLDLCPDCEFGHPDTTYAK 596

Qy 601 CGSGPMLTPRCLIDPYRLMHYPTVNTYTFKIRMYGVGEHRLTPACNFTRGDRCNLED 660
Db 597 CGSGPMLTPRCLIDPYRLMHYPTVNTYTFKIRMYGVGEHRLTPACNFTRGDRCNLED 656

Qy 661 RDRSLSPLSHSTTEMAILPGSYSDLPALSTGLHLHONIUDVOFMYGLSPALTKYVRW 720
Db 657 RDRSLSPLSHSTTEMAILPGSYSDLPALSTGLHLHONIUDVOFMYGLSPALTKYVRW 716

Qy 721 EFWVILFLILADARVACIIMWLILGQAEALAEKIVIIHAASAACNGFLYVIPVAVM 780
Db 717 EFWVILFLILADARVACIIMWLILGQAEALAEKIVIIHAASAACNGFLYVIPVAVM 776

Qy 781 YIKGRVPLATYSITGLWSFSLILALPOAAVAYDASVHGQIALLVITFTLTPGYK 840
Db 777 YIKGRVPLATYSITGLWSFSLILALPOAAVAYDASVHGQIALLVITFTLTPGYK 836

Qy 841 TILSRPLMWLCYLLTLGEMVQEMAPPMQVRCGRDGIITMAVAIFPGVVFDTTKMLAVL 900
Db 837 VILARILMWLQYFTTRAEDLHVMTIPPLNARSGRAIILLMCAVIBELFDITKILIALI 896

Qy 901 GPAYILKALTRVPEYVAHALBMCMTABHLAGRYVOMALALGRMTGTIYDHLTPM 960
Db 897 GPAYILKALTRVPEYVAHALBMCMTABHLAGRYVOMALALGRMTGTIYDHLTPM 956

Qy 961 SDMAASGLRDLAVAVEPIIFSPMEKKVIWGAETACDILHGLPVSARLIGREVLLGPD 1020
Db 956 SDMAASGLRDLAVAVEPIIFSPMEKKVIWGAETACDILHGLPVSARLIGREVLLGPD 1020

Db 957 RDMPRAGLRLALAVNEBVFVSDMETKITTWGADTLACGDIIILGLPVASARRKEIILLGPDAD 1016
 Qy 1021 GYTSKWSLAPITAYAOQTRGLGITVSMGTGDKTEOAGEIQVLSVTYQFSGTISG 1080
 Db 1017 SLBGRGLRLAPITAYAOQTRGLGCIITSLTGRDKNQVEGVVSTANQSLATCVNG 1076
 Qy 1081 VMTVTHGAGNKLTAGRGVLTQWYSSABEDLYGWSEPPPGTKSLPECTCGADVLYLVTRN 1140
 Db 1077 VCVTVVHAGSKTLAARPGPITQWYTVNDODLVGMPKPPGARSILPTCTCSSSLYLVTRH 1136
 Qy 1141 ADVIPARRRGDKGALISPRPLSTLKSGSGGPVLCPEGHAVGVFPAVCSRGVAKSIDF 1200
 Db 1137 ADVIPARRRGDSKSLSPRPVSTLKSGSGGPLCPGHAHVGLFPAVCTRGVAKAVDF 1196
 Qy 1201 PVETLDIVTRSPTEFSDNSTPRAVPQTVQVYLAHAPTSKSTKVPVAYAAQYKVLVLP 1260
 Db 1197 PVSEMETTMSPVYETDNSSPPAVPQSFVHAHLHAPTSKSTKVPVAYAAQYKVLVLP 1256
 Qy 1261 SVATTLGFGAYLSKAHGINPNIRGVTVTTGAPITTYTGKFLADGGCAGAYDIILICD 1320
 Db 1257 SVATTLGFGAYMSKAHGINPNIRGVTVTTGAPITTYTGKFLADGGCAGAYDIILICD 1316
 Qy 1321 ECHAVDSTTLIGITVLDQAEAGVRLTVLATAATPPGSVTPPHNTEEVALGQEGEIPFV 1380
 Db 1317 ECHVSTSTTLIGITVLDQAEAGVRLTVLATAATPPGSVTPPHNTEEVALSMTGELIPFV 1376
 Qy 1381 GRAIPLSYIKGRHLIFCHSKKKCDLAAALRGMLNSVAYYRGLDVSVIPTQGDVVVA 1440
 Db 1377 GKAIPIAIRGGRHLIFCHSKKKCDLAAALRGMLNSVAYYRGLDVSVIPTIGDVVVA 1436
 Qy 1441 TDLMTGVTGDFDSVIDCNNAVTVQVDFSLDPFTTTITQVYPODANVSRRRGRTGRGL 1500
 Db 1437 TDLMTGVTGDFDSVIDCNNAVTVQVDFSLDPFTTTITQVYPODANVSRRRGRTGRGR 1496
 Qy 1501 GIYRVYSTGRBASGMPDSVVLCECYDAGAAWELTPESETVRLRAYNTEGFLPVCDHLE 1560
 Db 1497 GIYRVYTPGERPSGMPDSVVLCECYDAGAAWELTPESETVRLRAYNTEGFLPVCDHLE 1556
 Qy 1561 FWEAVFTGLTHIDAHFLSOTKQSGENFAYLTAQVATVCAKAKAPPSMDVMWKLTRLK 1620
 Db 1557 FWEAVFTGLTHIDAHFLSOTKQSGENFAYLTAQVATVCAKAKAPPSMDVMWKLTRLK 1616
 Qy 1621 TLVGPFTLLRLSGVTVNEVTLTHPVTKYIATCMQADIEWNTSTWVLAGVTLAAVAYCLA 1680
 Db 1617 TLVGPFTLLRLSGVTVNEVTLTHPVTKYIATCMQADIEWNTSTWVLAGVTLAAVAYCLA 1676
 Qy 1681 TGCVCIIGRLHINORAVAVAPDKVLYEAPDEMECASRALIEBQIAEMLSKIQGL 1740
 Db 1677 TGCVCIIGRLHINORAVAVAPDKVLYEAPDEMECASRALIEBQIAEMLSKIQGL 1736
 Qy 1741 QOASKQADIOPTVQASMPKVEOFWAKHNNFISGIQYLAGLSTLPQNPVAVASMMAFSA 1800
 Db 1737 QOATKQAEAAAPVYESSKRALETFWAKHNNFISGIQYLAGLSTLPQNPVAVASMMAFSA 1796
 Qy 1801 LTPSLSTSTTLINLTLCGMLASQIAPRAGATGPVSGVLNGAVSGTGLGVLYDIIAGY 1860
 Db 1797 LTPSLSTSTTLINLTLCGMLASQIAPRAGATGPVSGVLNGAVSGTGLGVLYDIIAGY 1856
 Qy 1861 AGISGALVAFKINGEKSPMEDVNLPGILSPGALVGYICAILRRHVGPGEGAVQM 1920
 Db 1857 AGVAGALVAFKINGEKSPMEDVNLPGILSPGALVGYICAILRRHVGPGEGAVQM 1916
 Qy 1921 NRIIAFASRGNHVAPTHVYTESDASQRTQGLSTLTLRLHNMITECDPLPCGGSW 1980
 Db 1917 NRIIAFASRGNHVAPTHVYTESDASQRTQGLSTLTLRLHNMITECDPLPCGGSW 1976
 Qy 1981 LRDVMDVAVCTILTPFKXWLSKLEPKMPGLPVSCQGYGVGAWGTGIMTRPCGANIS 2040
 Db 1977 LRDVMDVAVCTILTPFKXWLSKLEPKMPGLPVSCQGYGVGAWGTGIMTRPCGANIS 2036
 Qy 2041 GNVLSGMRITGPTKCNMIMOGTEPINCYTEGQCVPPAPANFKVAILRVVASEVAYEATOH 2100
 Db 2037 GNVLSGMRITGPTKCNMIMOGTEPINCYTEGQCVPPAPANFKVAILRVVASEVAYEATOH 2096

Qy 2101 GSYHYITGLITDNLKVPQCLPSPPEFSWVGQIHRFAPTPKPFREDEVSFCVGLNSFVY 2160
 Db 2097 GDFHVTWGTITDNLKVPQCLPSPPEFSWVGQIHRFAPTPKPFREDEVSFCVGLNSFVY 2156
 Qy 2161 GSOQLCPDEPPDIDVMSMLTDPSSHITAEFAARLARGSPSPASSASQOLSPSLRATCT 2220
 Db 2157 GSOQLCPDEPPDIDVMSMLTDPSSHITAEFAARLARGSPSPASSASQOLSPSLRATCT 2216
 Qy 2221 THGKAVDVMVDFANLF---MGADVTRIESGSKVVVLSLDPMVEERSDLEPSISEWYL 2276
 Db 2217 THGKAVDVMVDFANLF---MGADVTRIESGSKVVVLSLDPMVEERSDLEPSISEWYL 2276
 Qy 2277 PKKREPPALPAMARPDVNPPLVESWKRDPDYQATVAGCALPPPKTPTPPPPRRRTVGLS 2336
 Db 2277 PKKREPPALPAMARPDVNPPLVESWKRDPDYQATVAGCALPPPKTPTPPPPRRRTVGLS 2336
 Qy 2337 EDSIGDALQOLAKSPGQPPSGDGLSTGAGAASSGQTP-PDELA-----LSETGIS 2390
 Db 2337 EDSIGDALQOLAKSPGQPPSGDGLSTGAGAASSGQTP-PDELA-----LSETGIS 2386
 Qy 2391 SMPPLEGELGDPDLPEQVEPQPPQGVAAAGSDSGSWSCTSEB-SDSVVCCSSWSWT 2449
 Db 2387 SMPPLEGELGDPDLPEQVEPQPPQGVAAAGSDSGSWSCTSEB-SDSVVCCSSWSWT 2426
 Qy 2450 GALITPCSEBEKLPINPLNSLRLYHNKVCYCTTKASLRAKKVTPDRMOVLSYDVS 2509
 Db 2427 GALITPCSEBEKLPINPLNSLRLYHNKVCYCTTKASLRAKKVTPDRMOVLSYDVS 2486
 Qy 2510 LKDIKLAASKVYARLLTMEACOLTPPHSABSKYFSGKAVSLSGRAVNHKSVMKDL 2569
 Db 2487 LKDIKLAASKVYARLLTMEACOLTPPHSABSKYFSGKAVSLSGRAVNHKSVMKDL 2546
 Qy 2570 EDSSETPIPTTINAKREVEFCVDPTEKGAARLIVPDLGVRVCEKMAUYDITOKLPQAVM 2629
 Db 2547 EDSSETPIPTTINAKREVEFCVDPTEKGAARLIVPDLGVRVCEKMAUYDITOKLPQAVM 2606
 Qy 2630 GASYGFQYSPAORVEFLKAMAEKKDPMGFSYDTRCFDSVTTERDITRTEESITYRACSLPE 2689
 Db 2607 GASYGFQYSPAORVEFLKAMAEKKDPMGFSYDTRCFDSVTTERDITRTEESITYRACSLPE 2666
 Qy 2690 EAHYTAHSLTERLYVGGPMFNSKQOTCGRRRASAAGVLTTSNGNITTCVRLAAKCKAG 2749
 Db 2667 EAHYTAHSLTERLYVGGPMFNSKQOTCGRRRASAAGVLTTSNGNITTCVRLAAKCKAG 2726
 Qy 2750 IIAPTMLVCGDLDVYISSESQTEEDERNLAFTEAMTRYSAAPGDPREYDLEHITSCS 2809
 Db 2727 IIAPTMLVCGDLDVYISSESQTEEDERNLAFTEAMTRYSAAPGDPREYDLEHITSCS 2786
 Qy 2810 SNVSVALGFGRRRYLTRDPTTIPARAAMEFTRHSPVNSWLGNTIOTAPTIWAMVLT 2869
 Db 2787 SNVSVALGFGRRRYLTRDPTTIPARAAMEFTRHSPVNSWLGNTIOTAPTIWAMVLT 2846
 Qy 2870 HFFSILIMADDTLDONLNEFMGAVSVSPDLPAIIBRLHGLDAPSLHTYTHELTRVVS 2929
 Db 2847 HFFSILIMADDTLDONLNEFMGAVSVSPDLPAIIBRLHGLDAPSLHTYTHELTRVVS 2906
 Qy 2930 ALRKLGAPELRAMKSRARARASLISRGGRAAVCGRYLFNMAVKTLLKLTPLPEARLIDL 2989
 Db 2907 ALRKLGAPELRAMKSRARARASLISRGGRAAVCGRYLFNMAVKTLLKLTPLPEARLIDL 2966
 Qy 2990 SSMFTVAGAGSDIYHSVSRAPRLILFGLLLPVGVGLFLPAR 3033
 Db 2967 SSMFTVAGAGSDIYHSVSRAPRLILFGLLLPVGVGLFLPAR 3010

RESULT 12
 US-08-904-686A-14
 ; Sequence 14, Application US/08904686A
 ; Patent No. 5998130
 ; GENERAL INFORMATION:
 ; APPLICANT: OKRAYAMA, Hiroto
 ; APPLICANT: FUKU, Isao


```

APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESS: Armstrong, Westernman, Hattori, Mclelland &
ADDRESS: Naughton
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/904,686A
FILING DATE: 01-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Mclelland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 900703G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 3010 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-904-686A-14

```

```

Query Match 74.3%; Score 11984.5; DB 2; Length 3010;
Best Local Similarity 71.5%; Pred. No. 0;
Matches 2175; Conservative 352; Mismatches 472; Indels 45; Gaps 8;

```

```

QY 1 MSTNPKRQKTKXNTNRPRDQVYFPGGQIVGVYLLPRGRPLGTRATRKTSERQPRG 60
DB 1 MSTNPKRQKTKXNTNRPRDQVYFPGGQIVGVYLLPRGRPLGTRATRKTSERQPRG 60
QY 61 RROP1PKDRSTGKSGKFGYPMPLXGNBGLGAGWMLSPRGSRPMSGPNPDRHRSRNYG 120
DB 61 RROP1PKDRSTGKSGKFGYPMPLXGNBGLGAGWMLSPRGSRPMSGPNPDRHRSRNYG 120
QY 121 KYIDTITCGPADIMGYIPVVGAPLGVAPALAHGAVLEDDGVNFATGNLPGCSFSIFLLA 180
DB 121 KYIDTITCGPADIMGYIPVVGAPLGVAPALAHGAVLEDDGVNFATGNLPGCSFSIFLLA 180

```

```

QY 181 LLSCTTPVSAABVKNIISGYWVNDCTNDSITWOLQAALVHPGCPCKEKNASQCMW 240
DB 181 LLSCTTPVSAABVKNIISGYWVNDCTNDSITWOLQAALVHPGCPCKEKNASQCMW 240
QY 241 PVSNNVAVORPGALTOGLRTHIDMVMSATLCSALYVGDLCGVMLAQMFTVSPQHMF 300
DB 241 PVSNNVAVORPGALTOGLRTHIDMVMSATLCSALYVGDLCGVMLAQMFTVSPQHMF 300
QY 301 VQDCNCSTYPTITGHRAMDMMMMSPTATMILAYAMVPEVIIDIGAHKGVNFGLA 360
DB 301 VQDCNCSTYPTITGHRAMDMMMMSPTATMILAYAMVPEVIIDIGAHKGVNFGLA 360
QY 361 YFSMOGAMKVVVILLIAGVDAFRTHTGVSAAQTGRITSLFDMGPRKIQOLVUNNGSM 420
DB 361 YFSMOGAMKVVVILLIAGVDAFRTHTGVSAAQTGRITSLFDMGPRKIQOLVUNNGSM 420
QY 421 HINRTALNCNDSLHTGFIASLFYTHSPNSSGCPERRMSACRS1EAFRVGALQYEDNVN 480
DB 421 HINRTALNCNDSLHTGFIASLFYTHSPNSSGCPERRMSACRS1EAFRVGALQYEDNVN 480
QY 481 PEDMRPYCWHYPPROCGVVSATVCGPVYCFTPSPVYVGTTRBLGAPTYTWGENETDVL 540
DB 481 PEDMRPYCWHYPPROCGVVSATVCGPVYCFTPSPVYVGTTRBLGAPTYTWGENETDVL 540
QY 541 LNSTRPGLSGWFGCTMMSSGYTKTCGAPPCRTADFNASTDILCPTDCFRKHDPDTTYL 600
DB 541 LNSTRPGLSGWFGCTMMSSGYTKTCGAPPCRTADFNASTDILCPTDCFRKHDPDTTYL 600
QY 601 CGSGPWLTPRCLIDYPRYLMHYPCVTNVTIPIKIRMYGVGEHRLTAACNFTGRDCNLED 660
DB 601 CGSGPWLTPRCLIDYPRYLMHYPCVTNVTIPIKIRMYGVGEHRLTAACNFTGRDCNLED 660
QY 661 RRSQSLPILHSTTEWALIPCSYSDLPALSTGLHLHONIYVQFMYGSLPALTXYRN 720
DB 661 RRSQSLPILHSTTEWALIPCSYSDLPALSTGLHLHONIYVQFMYGSLPALTXYRN 720
QY 721 EYVLLFLLIADARVACACIMMLIILGOAEALAEKVIILHAASAASNGFLYFVIFVAM 780
DB 721 EYVLLFLLIADARVACACIMMLIILGOAEALAEKVIILHAASAASNGFLYFVIFVAM 780
QY 781 YIKGRVPLATYSLTGLWSFSLILLALPOAAVAYDASVHGQIAGALLVMTITFTLPYK 840
DB 781 YIKGRVPLATYSLTGLWSFSLILLALPOAAVAYDASVHGQIAGALLVMTITFTLPYK 840
QY 841 TILSRFLMWLYLLTGLGAMVOEMAPPMOVRGRCGCIYAAVAIFIPGVVFDITKMLANV 900
DB 841 TILSRFLMWLYLLTGLGAMVOEMAPPMOVRGRCGCIYAAVAIFIPGVVFDITKMLANV 900
QY 901 GPAYLLKGLTVPRYFVAHAALLRMCTMARHLAAGRYVOMALLAGRWGTIYDHLTEM 960
DB 901 GPAYLLKGLTVPRYFVAHAALLRMCTMARHLAAGRYVOMALLAGRWGTIYDHLTEM 960
QY 961 SDMAASGLRDLAFAVEPIIFSPMEKKVIYMGAEATAACGDIILHGLPVSARGLREVLLGPD 1020
DB 961 SDMAASGLRDLAFAVEPIIFSPMEKKVIYMGAEATAACGDIILHGLPVSARGLREVLLGPD 1020
QY 1021 GYTSKGSMLADITYAAOQTRGLCTIVSMGTGRKTEQAGEIYQLSTVYTSFGLTGISG 1080
DB 1021 GYTSKGSMLADITYAAOQTRGLCTIVSMGTGRKTEQAGEIYQLSTVYTSFGLTGISG 1080
QY 1081 VLMATVYHAGNKTILAGSGRPVOMYSSABGDLVGMSPRGCTSLPCTCGAADVLYVTN 1140
DB 1081 VLMATVYHAGNKTILAGSGRPVOMYSSABGDLVGMSPRGCTSLPCTCGAADVLYVTN 1140
QY 1141 ADVIPARRRGDKRGAALLSPRLSTLKSSGGVLCPRGHAVGVFPAAVCSRSVAKSIDFI 1200
DB 1141 ADVIPARRRGDKRGAALLSPRLSTLKSSGGVLCPRGHAVGVFPAAVCSRSVAKSIDFI 1200
QY 1201 PVEITDIYTRSPEDNSTPPAVPOTYOVGYIAHATGSGSKTKVVAIAAOGYKTLVYLP 1260
DB 1201 PVEITDIYTRSPEDNSTPPAVPOTYOVGYIAHATGSGSKTKVVAIAAOGYKTLVYLP 1260
QY 1261 PVESMETTMRSPVFTDSSPPAVPOSFOVALHAAPTSGSKTKVPAAYAAOQYKVLVLP 1320
DB 1261 PVESMETTMRSPVFTDSSPPAVPOSFOVALHAAPTSGSKTKVPAAYAAOQYKVLVLP 1320

```

QY 1261 SVAATLGFAGVLSAAGINPNIRTVRTTGTABITYSTYCKFLADGGCAGAYDIIICD 1320
 DB 1257 SVAATLGFAGVLSAAGINPNIRTVRTTGTABITYSTYCKFLADGGCAGAYDIIICD 1316
 QY 1321 ECHAVDSTTLTGCTVLDQAEAGVRLVLTATAPPSVYTPPHNIEEVALGQEBEIPFY 1380
 DB 1317 ECHSTDTSTTLTGCTVLDQAEAGVRLVLTATAPPSVYTPPHNIEEVALGQEBEIPFY 1376
 QY 1381 GRALPLSYIKGGRHLIFCHSKKKCDLAAALRGWGLSVAYYRGLDVSVPFGQDVVVVA 1440
 DB 1377 GKALPIEALRGGRHLIFCHSKKKCDLAAALRGWGLSVAYYRGLDVSVPFGQDVVVVA 1436
 QY 1441 TDALMTGTGDPDSVIDCNNAVTVQVDBSLDPTTTTQTVYQDASRSQRRGTGRGL 1500
 DB 1437 TDALMTGTGDPDSVIDCNNAVTVQVDBSLDPTTTTQTVYQDASRSQRRGTGRGL 1496
 QY 1501 GIYVYVSTGERASGMPFSVULCECYDAGAAWELTPBETVRLAAYNTGTGLPVCODHLE 1560
 DB 1497 GIYVYVSTGERASGMPFSVULCECYDAGAAWELTPBETVRLAAYNTGTGLPVCODHLE 1556
 QY 1561 FMEAVFTGLTHIDAHFLSQTKQSGENFAVLTAYOATVCARAKAPPSPMDVMMKCLTRLPK 1620
 DB 1557 FMEAVFTGLTHIDAHFLSQTKQSGENFAVLTAYOATVCARAKAPPSPMDVMMKCLTRLPK 1616
 QY 1621 TLVGPPTLLYRLSGVTNEVTLTHPVYKIATQWQADLEWMTSTWVLAVGAAYAAVYCLA 1680
 DB 1617 TLHGPPTLLYRLSGVTNEVTLTHPVYKIATQWQADLEWMTSTWVLAVGAAYAAVYCLA 1676
 QY 1681 TGCCCTIGRLHINORAVAPDKEVLYAEFDEMECASRALIEGQRIAEMLSKIOGLL 1740
 DB 1677 TGSAVVIYGRILTLISGRPALVPDRRELYQFDEMECASRALIEGQRIAEMLSKIOGLL 1736
 QY 1741 QOASQKQADIDOPTQASWPKVKEQOFMAKHMNFISIGIYLAGLSTLPQNPAAVMAWMAASAA 1800
 DB 1737 QOATKQKQAAAPVAVESKRALETFTMAKHMNFISIGIYLAGLSTLPQNPAAVMAWMAASAA 1796
 QY 1801 LTPSLSTSTTLNLILGGLWLSQIAPAPGATGVVSGVGLVGAAGSILGKLVLDIILAGYG 1860
 DB 1797 LTPSLSTSTTLNLILGGLWLSQIAPAPGATGVVSGVGLVGAAGSILGKLVLDIILAGYG 1856
 QY 1861 AGISGALVAFKIMSGEKPMEDVNLPLGISGALVAVGVCAILIRRHVGPGGAQVM 1920
 DB 1857 AGVAGALVAFKIMSGEKPMEDVNLPLGISGALVAVGVCAILIRRHVGPGGAQVM 1916
 QY 1921 NRLIAFASRGNHVAPTHVYTESDASQRYTQLGSLITSLRLRHNMITECCPIPCGSGW 1980
 DB 1917 NRLIAFASRGNHVAPTHVYTESDASQRYTQLGSLITSLRLRHNMITECCPIPCGSGW 1976
 QY 1981 LRVDVAVCTIILDPKMWLTSKLPKMGELPFVSCQKYGKVAAGTGMITTRCPGANIS 2040
 DB 1977 LRVDVAVCTIILDPKMWLTSKLPKMGELPFVSCQKYGKVAAGTGMITTRCPGANIS 2036
 QY 2041 GNVVLGSMRITGPTCKMIMNGTTPINCYTEGOCVPPAPAPFKAIIRVVAASEVATQH 2100
 DB 2037 GNVVLGSMRITGPTCKMIMNGTTPINCYTEGOCVPPAPAPFKAIIRVVAASEVATQH 2096
 QY 2101 GSYHYITGLTNDNLKVPQCLSPPEFFSVVDGVQIHRFAPTRKPPFRDEVSFCVGLNFVV 2160
 DB 2097 GDFHYVGMITDNLKVPQCLSPPEFFSVVDGVQIHRFAPTRKPPFRDEVSFCVGLNFVV 2156
 QY 2161 GSQLPCEBEPDVLMSMLTDPSHITAEARLARGSPSEASSASQLSAPSLRATCT 2220
 DB 2157 GSQLPCEBEPDVLMSMLTDPSHITAEARLARGSPSEASSASQLSAPSLRATCT 2216
 QY 2221 THGKAVDVMADANLF----NGDVTREISGSKVVLDSLDPMVEESDLEPSISRYML 2276
 DB 2217 THGKAVDVMADANLF----NGDVTREISGSKVVLDSLDPMVEESDLEPSISRYML 2276
 QY 2277 PKKFPFALPAMAPADVNPPLVESWKRBDYQATVAGCALPPPKRTPPPRRRRRTYGLS 2336
 DB 2277 KSKKFPFALPAMAPADVNPPLVESWKRBDYQATVAGCALPPPKRTPPPRRRRRTYGLS 2336
 QY 2337 EDSIGDALQOLAIRSGQPPSGDSGLSTGAGAADSGSQTP-PDELA-----LSETGSIS 2390

DB 2337 ESSVSALAEIATKTFG-----SSSSAVDSGTATALPDQASDGDGKGSVSES 2386
 QY 2391 SMPLEGEIGDDBLDEPOVEPPPPQGVAAAGSGSGMSSTSEE-DSVVCSSMSYMT 2449
 DB 2387 SMPLEGEIGDDBL-----SDGSSTVSEASEBVCCSSSYMT 2426
 QY 2450 GALITPCSPBEKLPINPLSNSILRYHNKVCYCTTKASALRAKAVTFPRMQVLDSYDSV 2509
 DB 2427 GALITPCAESEKLPINPLSNSILRHHNMVAVTTSRAGLRQKVTLPRLQLDHDYRV 2486
 QY 2510 LKDIKLAASKVTRALLTMEBAQULPPHSABKYGFGAKVNSLSCRAVNHKSYWKDL 2569
 DB 2487 LKEMKAKSTVAKKLISVEACKLTPPHSAKSKFGYGAADVNLSSKAVNHLSYWKDL 2546
 QY 2570 EDESEPIPTTMAKNEVFCVDPPTKGGKKAARLIYVPDGVRCEKALYDIOXKLPOAVM 2629
 DB 2547 EDVTPIDITTIMAKNEVFCVDPPTKGGKKAARLIYVPDGVRCEKALYDIOXKLPOAVM 2606
 QY 2630 GASYGFQSPAOVRVELKAKMAEKDPMGFSYDTCFDPSTVERDIRTEESYIRACSLPE 2689
 DB 2607 GSSYGFQSPAOVRVELKAKMAEKDPMGFSYDTCFDPSTVERDIRTEESYIRACSLPE 2666
 QY 2690 EAHTAHSLTERLYYGGEMFNKSGTCGRCRASGVLTTSKGNITTCYKALACKAAG 2749
 DB 2667 EAHTAHSLTERLYYGGEMFNKSGTCGRCRASGVLTTSKGNITTCYKALACKAAG 2726
 QY 2750 IIAPTLVGDDLVVYSESQTEDEERNLRAETBAMTRYSAPGDPREVELEYITSCS 2809
 DB 2727 IODCTLVAGDDLVVYSESQTEDEERNLRAETBAMTRYSAPGDPREVELEYITSCS 2786
 QY 2810 SNVSVALGPOGRRRYLTREDPTTPIARAAMEFTRHSPVNSWLGNTIOVAPTIMABVLM 2869
 DB 2787 SNVSVALGPOGRRRYLTREDPTTPIARAAMEFTRHSPVNSWLGNTIOVAPTIMABVLM 2846
 QY 2870 HFFSILMAODTIDONLNFEMGAVSVSPLDPAIIEHLHGDAISLHTYTHELTRVAS 2929
 DB 2847 HFFSILMAODTIDONLNFEMGAVSVSPLDPAIIEHLHGDAISLHTYTHELTRVAS 2906
 QY 2930 ALRKLGAAPPLRAWKSARAVRASLISRGRAAVCGRYLENMAVKTLPKTPPEARLDTL 2989
 DB 2907 ALRKLGAAPPLRAWKSARAVRASLISRGRAAVCGRYLENMAVKTLPKTPPEARLDTL 2966
 QY 2990 SSWFTVAGAGGDIYHSVSRARPRLLFGILLFVGVLELIPBAR 3033
 DB 2967 SSWFTVAGAGGDIYHSVSRARPRLLFGILLFVGVLELIPBAR 3010

RESULT 13
 US-09-315-850-2
 / Sequence 2, Application US/09315850
 / Patent No. 6217872
 / GENERAL INFORMATION:
 / APPLICANT: OKAYAMA, Hiroto
 / APPLICANT: FURE, Iseao
 / APPLICANT: MORI, Chisato
 / APPLICANT: TAKAMIZAWA, Akahisa
 / APPLICANT: YOSHIDA, Iwao
 / TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
 / NUMBER OF SEQUENCES: 50
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Armstrong, Westerman, Hattori, McIreland &
 / STREET: 1725 K St. N.W. Suite 1000
 / CITY: Washington
 / STATE: D.C.
 / COUNTRY: U.S.A.
 / ZIP: 20006
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0

```

SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/315,850
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/904,686
FILING DATE: 01-AUG-1997
APPLICATION NUMBER: US 08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-1-67466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-3-05605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: McLeeland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 900703G
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3010 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-315-850-2

Query Match      74.3%; Score 11984.5; DB 3; Length 3010;
Best Local Similarity 71.5%; Pred. No. 0;
Matches 2175; Conservative 352; Mismatches 472; Indels 45; Gaps 8;

```

```

421 HINRTALNCNDSLHNGFIASLFYTHSPNNSGGCBERMSACRSIEAFVWGALQYEDNVTN 480
421 HINRTALNCNDSLQGFLLALFYTHSPNNSGCCBERMAQCRITIDKPOGSGPIITYAE--SS 478
481 PEDMEPYCWHYPPROGVSAKTVGCVYCTPSPVWVGTTRDLGAPITYWGENETDVL 540
479 RDDQRYCWHYPPROGTVPAVEGCVYCTPSPVWVGTTRDLGAPITYWGENETDVL 538
541 LNSTRPPLGSGWCTGCTWMSGGYTKCGAPPCGTRADFNASTDLLCPTDFFRKHPDTTYK 600
539 LNSTRPPOGNWPGCTWMSNGTGTCTGCGPCNIGVGN--NLTCPTDFFRKHPDTTYK 596
601 CGSGPWLTPRCCLIDPYRLMHYPCYVNYTIFPIRMVGVGEHRLTAACNFTGDRCNLMD 660
597 CGSGPWLTPRCWVDPIYRLMHYPCYVNYTIFVRYMVGVEHRLNNAACWHTGERCDLMD 656
661 RDRSOLSPLNHSTTEWAILPCSYSDLPALSTGLHLHONIVDVQFMYGSLPALTKYIVM 720
657 RDRPELSPLLSTTEWQVLPFGFTLLPALSTGLHLHONIVDVQYLGISAVSFAIKM 716
721 EYVILLFLILADARVCACIMMLILAOEAALLENLVILNASVAGAHGILSELPFCAM 776
781 YIKGRVPLATYSILGLMSFSLLLALPOAAVADASVHGOI GAULLVMTLFTLPGYK 840
777 YIKGRVPLATYALGVWPLILLLALPPRAIAMDREMAASCAGAVFGLVLTLSPIYK 836
841 TLISREPLMWLCYLLTLGEMAVQEWAPPMQVRGDRDGIWAVALFYPGVVFDTITKMLAVL 900
837 VFLARLIWLOQFTTRAEADLHWMIPLNARSGRDAIILLMCAVHDELFDITKLLIAL 896
901 GPAYLLKALTVPPYFVRAHALLRMCYMARHLAGRYVOMALLALGRMTGTIYIDHLTPM 960
897 GELMVLQAGITVFPYFVRAQGLIHAICMLVRKVAAGHYVOMAPMKGALGTIYIYNHLEPL 956
961 SDMAASGLPDLAVAVEPIIFSPMEKKVYVWGETAACGDIILHGLPVSARLSREVLPGPD 1020
957 RDMPRAGLIDLAVAVEPVSDEMETKITWGDYTAACGDIILGLPVSARGKEILILGPD 1016
1021 GYTSKWSLILAPITVAOQTRGLIGTIVVSMGRDTEQAGEIOVLSTVTOGFLGTSISG 1080
1017 SLEGRGLRLALITVAISOQTRGLIGTISLGRDNQVEGVEVQVSTATQGFALTCVNG 1076
1081 VLMATVHAGNTTLAGSRGPVTOMYSAEGDVGWSPPTGTSLSBCTCGAVDLYLVTN 1140
1077 VCMATVHAGSKTTLAPKGPITOMYTNVDQDLVGMPPKPGARSILPTCTGSSDLYLVTH 1136
1141 ADVIPARRRGDKRGLLSRPLSTLKSSGSGVULCPRGHAGVVFRAVCSRGVAAISDPI 1200
1137 ADVIPARRRGDSRGLSPRPVSYLKGSSGGVULCPFGHAGVIFRAVCTTRGVAAKAVDV 1196
1201 PVETLIDVRSPTFSNDSTPPAVPQTYQVGYLHAPTSKSTKVPYAVAAQGYKVLVLP 1260
1197 PVESMETTRRSVPFLDNDSPRAVPQSFOYAHHAFTGSKSTKVPYAVAAQGYKVLVLP 1256
1261 SYAATLGFAGVYSKKAHINPNIRTVGRTVTTGAPITYSYGRFLADGGCAGAYDITCD 1320
1257 SYAATLGFAGVYSKKAHIDPNIRTVGRTVTTGAPITYSYGRFLADGGCAGAYDITCD 1316
1321 ECHAVDSTIILIGITVLDQAEAGVRLTVLATPSPGSTTPHPNIEEVALGOEGEIPRY 1380
1317 ECHVSTDSITLIGITVLDQAEAGVRLTVLATPSPGSTTPHPNIEEVALSNTGEIPRY 1376
1381 GRAIPLSYIKGGRHLIFCHSKKKCDELAALRGMGINSVAYVYRGDVSIVPTQGDVVVA 1440
1377 GKAIPIEALRGRHLIFCHSKKKCDELAALRGMGINSVAYVYRGDVSIVPTQGDVVVA 1436
1441 TDAIATGYGDDSDVIDCNVAVTQVVDPSLDPFTITITQIVQDAVSRQGRGRTRGL 1500
1437 TDALMTGYGDDSDVIDCNVTQVVDPSLDPFTITITQIVQDAVSRQGRGRTRGR 1496

```

Qy	1501	GIYVYVSGEASGAMPSSVVLCECYDGAAYELTPSESTYRLAAYNTPELPCOODHLE	1560
Db	1497	GIYFVYVPEGBPSGMPDSSVLCCECYDGCAYELTPAETSYRLAAYNTPELPCOODHLE	1556
Qy	1561	FWEAVFVGLTIDIAHFLSQTQSGENFAYLTAQYAQYCARAKAPPPSWDMKKCLTILKP	1620
Db	1557	FWEVSVFGLTHIDAHFLSQTQADGNFYLVAQYATVABARQAPPPSWDMKKCLILRKP	1616
Qy	1621	TLVPEPTLYRLSGSYNTEVTLTHPYTKIATCMQADLEVMSTSTWYLAGCVLAAYACLA	1680
Db	1617	TLHPTPTLYRLGAVONEVTLTHPTIKYIMCMSADLEVMSTWYLAGCVLAALAAACLT	1676
Qy	1681	TGCVCITIGRLINORAVAPDKEVLYEAFDEMEBACSAALIEBGOIRAEMLKSGIOGLL	1740
Db	1677	TGSVYIVGRITILSGRPALVPDRRELYOFEFDEMEBACSLPIYBQMQLAEQFKQKAGLL	1736
Qy	1741	QOASQKOADLOPTVOASWPKVEOPFAKIMWNFISGITYLAGLSTLPGNPAAVSMASAA	1800
Db	1737	QTATKQKAAAPVYVESKWRALETFAKIMWNFISGITYLAGLSTLPGNPALIASMAPTAS	1796
Qy	1801	LTPSLUSTITLLNLIGGMLASQIAPPAAGATGFWVSGLVGAVSGIGKXVLDILAGY	1860
Db	1797	ITSPLTQSTLLFNILGVMVAQIAPPAASAFVAGAGIAGAAGVSGIGKXVLDILAGY	1856
Qy	1861	AGISGALVAFKIMSEKSEMDVYNLPGIILSPGLVGVYCAALIRPHNPGGAVOM	1920
Db	1857	AGVAGALVAFVMSGEMPTEDLVNLPAILISPGALVGVVCAALIRRHVPGEGAQVM	1916
Qy	1921	NRLIAFASRGNHVAPTHVYTESDASORYTOLGSLTITSLRLRLNWTTEOCPIPCQGSW	1980
Db	1917	NRLIAFASRGNHVBSPTHVYTESDAAARTOTIOLSLTITQLKRLHQMINEOCSTPCSSW	1976
Qy	1981	LRDVMDVCTIITLDPKMLTSLKLPKMGDLFPVSCQKGYKGVAMGTGIMTTRCPGANTIS	2040
Db	1977	LRDVMDVCTIYTLDPKTMLOSKLPQLRPGVFPFSCQSGYKGVMRGDGIMQTTCCPAQIT	2036
Qy	2041	GNVNLGSRIRIGPRTCKMNINGTPTPINCYTGECQYPRAPFKVAILMPVAASYAEYTG	2100
Db	2037	GHVNGKSNRIVGPKTCSMTMHGTFFINAYTTGPPSPAPNYSRLAMVAAEEVYEVRV	2096
Qy	2101	GSYHYITGLTNDNLKPCQLSPSEFFSWVDQOIHFRFPTKPPFRDVSFCVGLNSFVY	2160
Db	2097	GDFFHYVGMTIDNVKPCQVAPPEFFSVDGRLHRYAPACRPLIREVYITQVGLNQYLV	2156
Qy	2161	GSQLPCEDEPTDLYMSMLTDPSHITAEATARLARLARGSPPEBASSASQLSAPSLRATCT	2220
Db	2157	GSQLPCEBEPTDVAULTSMLTDPSHITAEATARRLARLARGSPPLASSASQLSAPSLKATCT	2216
Qy	2221	THGKAYDVWDVANLF---MGDVTREESGSKYVYLDSDPMWEERSDLEPSIPESEVL	2276
Db	2217	THHVSPPADLIEANLIMRQEMGNTTRVESENKVVVLDSFPLRAEBDEBEREVSAPAEILR	2276
Qy	2277	PKKXFPPLPAMARPDYNPPLVESWKRBDYOPATAGCALPPPKRTPPPRRRRYVLS	2336
Db	2277	KSXKFPFPAAMPIMARPDYNPPLLESKXDYVPPVHGCPLPPIKAPPIPPRRKRITVLT	2336
Qy	2337	EDSIGDALQOLAIKSFQPPSPGDSGLSTGAGADSGQTP-PDELA-----LSBTGSI	2390
Db	2337	ESSVSSALAEATKTFG-----SSEBSADVSGTATLAPQASDGDGKSGVBEYS	2386
Qy	2391	SMPLEGLGDPDLEPEBOVEPOPPOQGVAAAPGSDSGSWSTCSBE-DSVYCCSMYSWT	2449
Db	2387	SMPLEGGPGBDL-----SDGWSVTSVEBASDVVCCSMYSWT	2446
Qy	2450	GALITPCGPEEBEKLPINFLSNSILRYHNKYCTTTKASLBAKKYTPRMQVLYDSYOSV	2509
Db	2427	GALITPCAEESKPIINALSNSLLEHNHMYVATTSRSGLRQKKTFFRLDYLDLHDYDV	2486
Qy	2510	LKDIATLASKYTARLITMEBAQCLPPPSASQYFGCAKEAYRSLSGRAYNHIKSYWKDIL	2569
Db	2487	LKENAKAKASYKAKLLSYBEAKCLTPPHSASKREYGAQDVRNLSKAYNHIHSWKDIL	2546
Qy	2570	EDSETPIPTTMAKNEVCPDPTGGKKAARLIYVDIGVRCERKALYDITOKLPOAVM	2629

```

Db 2547 EDVTPTDITIMAKNEKFCVQPEKGRKPRLLVFPDLGVRVCEKALVYVSTLPPVVM 2606
Qy 2630 GASVGFQYSPQAVVEFLVNTWKSKNPMGFSYDTRCFDSVTVERDIRTEESIYRACSLPE 2689
Db 2607 GSSVGFQYSPQAVVEFLVNTWKSKNPMGFSYDTRCFDSVTVERDIRTEESIYQCCDLAP 2666
Qy 2690 EARTAIHSLTERLYVGGPMNKSQGTGGRCAASGLVLTSMGNTICVYKALAAACGAAG 2749
Db 2667 EAQQAISLTERLYIGGPLTNSKGQNGYRCRCAAGVLTSSCGNTLTVCYKASACRAAK 2726
Qy 2750 IIAPTMLVCGDDVIVISESGTEDEBNRLAFTAMTRYSAAPGDPREPYDELITRSCS 2809
Db 2727 LQDCITMLVNGDDVIVICESAGTQEDASLRFTEAMTRYSAAPGDPREPYDELITRSCS 2786
Qy 2810 SNVSVALPGQRRRRYYLTRDPTTPIAARAEVYRHSFVNSLGNIIQYAPTIAWVLMY 2869
Db 2787 SNVSVAHADGSKRRYYLTRDPTTPIAARAEVYRHSFVNSLGNIIYAPLTIAWVLMY 2846
Qy 2870 HFSIILMAQOTLDONLNFEMYGAVSVSPDLDAITRLHGLDAFSLHTTTPHLLTVAS 2929
Db 2847 HFSIILMAQOLEKALDCCQYGCACYSLEPLDQLIERTLHGLSAFSLHSSYSGEINVAS 2906
Qy 2930 ALRDLGAPPLRAWKSRAVARASLISRGRAAVCGRYLFPMNAVYTKLKLPLPEARLLDL 2989
Db 2907 CLRDLGAPPLRAWKSRAVARASLISRGRAAVCGRYLFPMNAVYTKLKLPLPEARLLDL 2966
Qy 2990 SSWETVAGAGGDIYHSVSRARPLLIFGLLLFVGVGLFLLPAR 3033
Db 2967 SGWVNAVYSGDIYHLSRARPRFMCLLLLSGVGYILLPRN 3010

RESULT 14
US-09-315-850-14
: Sequence 14, Application US/09315850
: Patent No. 6217872
: GENERAL INFORMATION:
: APPLICANT: OKAYAMA, Hiroto
: APPLICANT: FUKU, Isao
: APPLICANT: MORI, Chisato
: APPLICANT: TAKAMIZAWA, Akahisa
: APPLICANT: YOSHIDA, Iwao
: TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
: TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
: NUMBER OF SEQUENCES: 50
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Armstrong, Westernman, Hattori, Mclelland &
: ADDRESSEE: Naughton
: STREET: 1725 K St. N.W. Suite 1000
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20006
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
: SOFTWARE: ASCII
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/315,850
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/904,686
: FILING DATE: 01-AUG-1997
: APPLICATION NUMBER: US 08/324,977
: FILING DATE: 18-OCT-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 2-167466
: FILING DATE: 25-JUN-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 2-230921
: FILING DATE: 31-AUG-1990
: PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: McLeand, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 900703G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 3010 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-315-850-14

```

```

Query Match      74.3%; Score 11984.5; DB 3; Length 3010;
Best Local Similarity 71.5%; Pred. No. 0;
Matches 2175; Conservative 352; Mismatches 472; Indels 45; Gaps 8;

```

```

QY 1 MSTNPKRQKTKRNTRRRPODVKFPGGGQIVGGVYLPRRGPRLGYRATRKTSERQPRG 60
DB 1 MSTNPKRQKTKRNTRRRPODVKFPGGGQIVGGVYLPRRGPRLGYRATRKTSERQPRG 60
QY 61 RROPPIKDRRSTGSKNGKQGYMPPLYGNEGLGAGWMLSPRSGRPSGWPNDPRHRNRYG 120
DB 61 RROPPIKDRRSTGSKNGKQGYMPPLYGNEGLGAGWMLSPRSGRPSGWPNDPRHRNRYG 120
QY 121 KYIDLTGCPADLMGYIPVVGAPLGGVABALAHGVNLEDDGVNFAQNLPGCSFSIFLLA 180
DB 121 KYIDLTGCPADLMGYIPVVGAPLGGVABALAHGVNLEDDGVNFAQNLPGCSFSIFLLA 180
QY 121 KYIDLTGCPADLMGYIPVVGAPLGGVABALAHGVNLEDDGVNFAQNLPGCSFSIFLLA 180
DB 121 KYIDLTGCPADLMGYIPVVGAPLGGVABALAHGVNLEDDGVNFAQNLPGCSFSIFLLA 180
QY 181 LLSCTITPVSAAEVKNISTGYMTNTDCTNDSITWQAAVLAHVPGVCPEKYNASQCM 240
DB 181 LLSCTITPVSAAEVKNISTGYMTNTDCTNDSITWQAAVLAHVPGVCPEKYNASQCM 240
QY 181 LLSCTITPVSAAEVKNISTGYMTNTDCTNDSITWQAAVLAHVPGVCPEKYNASQCM 240
DB 181 LLSCTITPVSAAEVKNISTGYMTNTDCTNDSITWQAAVLAHVPGVCPEKYNASQCM 240
QY 241 PVSNNVAVQRPGLTQGLRTHIDMVVMSATLCSALYVGDLCGGWMLAQMFIYSPQHMF 300
DB 241 PVSNNVAVQRPGLTQGLRTHIDMVVMSATLCSALYVGDLCGGWMLAQMFIYSPQHMF 300
QY 241 ALTPTLAANVTIPTTIRRHVDLIGAAAFCSAMTVGDLCSGVLVSQDLFTFSRRHYT 300
DB 241 ALTPTLAANVTIPTTIRRHVDLIGAAAFCSAMTVGDLCSGVLVSQDLFTFSRRHYT 300
QY 301 VDDCNCSITPGTITGRMAMDMMMNSPTATMTLAVAMRVEVITIDISGAHWGNFGLA 360
DB 301 VDDCNCSITPGTITGRMAMDMMMNSPTATMTLAVAMRVEVITIDISGAHWGNFGLA 360
QY 301 LDDCNCSITPGHVSGRHAMDMMNSPTTALVVSQDLRIPOAVDMVAGAHGVLAGLA 360
DB 301 LDDCNCSITPGHVSGRHAMDMMNSPTTALVVSQDLRIPOAVDMVAGAHGVLAGLA 360
QY 361 YFSMGAQAKVVVYLILLAGVDARHTVVGSAQAOTGRLTSLFDMPROKIQOLVNTNGS 420
DB 361 YFSMGAQAKVVVYLILLAGVDARHTVVGSAQAOTGRLTSLFDMPROKIQOLVNTNGS 420
QY 361 YFSMGAQAKVVVYLILLAGVDARHTVVGSAQAOTGRLTSLFDMPROKIQOLVNTNGS 420
DB 361 YFSMGAQAKVVVYLILLAGVDARHTVVGSAQAOTGRLTSLFDMPROKIQOLVNTNGS 420
QY 421 HINRTALNCNDSLHTGFIASLFTYHSFNSSGCEBMSACRSIEAFRVGALQYEDNVN 480
DB 421 HINRTALNCNDSLHTGFIASLFTYHSFNSSGCEBMSACRSIEAFRVGALQYEDNVN 480
QY 421 HINRTALNCNDSLHTGFIASLFTYHSFNSSGCEBMSACRSIEAFRVGALQYEDNVN 480
DB 421 HINRTALNCNDSLHTGFIASLFTYHSFNSSGCEBMSACRSIEAFRVGALQYEDNVN 480
QY 481 PEDMPPYCMHYPPROGCVSAKTVGCVVCFPTSPVYVGTTRDLAPYTTGNEEDVPL 540
DB 481 PEDMPPYCMHYPPROGCVSAKTVGCVVCFPTSPVYVGTTRDLAPYTTGNEEDVPL 540
QY 479 RSDQRYCMHYPPROGCVSAKTVGCVVCFPTSPVYVGTTRDLAPYTTGNEEDVPL 538
DB 479 RSDQRYCMHYPPROGCVSAKTVGCVVCFPTSPVYVGTTRDLAPYTTGNEEDVPL 538
QY 541 LNSTRPPLSGWFCGTMMNSGGYTKTGAPPCRTADFNASTDLCTPDGFRKHPDTTYLK 600
DB 541 LNSTRPPLSGWFCGTMMNSGGYTKTGAPPCRTADFNASTDLCTPDGFRKHPDTTYLK 600
QY 539 LINTREPQGNMFCGTMMNSGGYTKTGAPPCRTADFNASTDLCTPDGFRKHPDTTYLK 596
DB 539 LINTREPQGNMFCGTMMNSGGYTKTGAPPCRTADFNASTDLCTPDGFRKHPDTTYLK 596
QY 601 CGSGPMLTRCLIDYYRLMHTPTCTYNTTIFKIRMTYGVGEHRLTACNFTRGDCNLED 660
DB 601 CGSGPMLTRCLIDYYRLMHTPTCTYNTTIFKIRMTYGVGEHRLTACNFTRGDCNLED 660
QY 597 CGSGPMLTRCLIDYYRLMHTPTCTYNTTIFKIRMTYGVGEHRLTACNFTRGDCNLED 656
DB 597 CGSGPMLTRCLIDYYRLMHTPTCTYNTTIFKIRMTYGVGEHRLTACNFTRGDCNLED 656

```

```

QY 661 RDRSOLSPILHSTTEWAILPCSYSDLPALSTGLIHLHONIVDQVPMYGLSPALTKYIVR 720
DB 661 RDRSOLSPILHSTTEWAILPCSYSDLPALSTGLIHLHONIVDQVPMYGLSPALTKYIVR 720
QY 657 RDRBELSPILHSTTEWAILPCSYSDLPALSTGLIHLHONIVDQVPMYGLSPALTKYIVR 716
DB 657 RDRBELSPILHSTTEWAILPCSYSDLPALSTGLIHLHONIVDQVPMYGLSPALTKYIVR 716
QY 721 EWLILFLILADARVACACMLMLILGOAEAALEKYLILHAASNASNGFLPYVIFVVAAM 780
DB 721 EWLILFLILADARVACACMLMLILGOAEAALEKYLILHAASNASNGFLPYVIFVVAAM 780
QY 717 EYLILFLILADARVACACMLMLILGOAEAALEKYLILHAASNASNGFLPYVIFVVAAM 776
DB 717 EYLILFLILADARVACACMLMLILGOAEAALEKYLILHAASNASNGFLPYVIFVVAAM 776
QY 781 YIKGRVPLATYSLTSLMSFSLILALPOOAYAYDASVHGQIGALLVMTLFTLPYK 840
DB 781 YIKGRVPLATYSLTSLMSFSLILALPOOAYAYDASVHGQIGALLVMTLFTLPYK 840
QY 777 YIKGRVPLATYSLTSLMSFSLILALPOOAYAYDASVHGQIGALLVMTLFTLPYK 836
DB 777 YIKGRVPLATYSLTSLMSFSLILALPOOAYAYDASVHGQIGALLVMTLFTLPYK 836
QY 841 TLLSRFLMVCYLITLGEAMVOEMAPPMQVRGDRGDIINAVALIFGYGVPDITKMLAV 900
DB 841 TLLSRFLMVCYLITLGEAMVOEMAPPMQVRGDRGDIINAVALIFGYGVPDITKMLAV 900
QY 837 VFLARLIMWLYQFTTRAEADLHVLPPLNARGRDAILLMCAVHPELLFDITLILAIL 896
DB 837 VFLARLIMWLYQFTTRAEADLHVLPPLNARGRDAILLMCAVHPELLFDITLILAIL 896
QY 901 GPAYLLKALTRVPEVFRALHALLRMCTMARHLAGRYVQVMAALALGRMTGYIYDHLTPM 960
DB 901 GPAYLLKALTRVPEVFRALHALLRMCTMARHLAGRYVQVMAALALGRMTGYIYDHLTPM 960
QY 897 GPLMVLQAGITTRVPEVFRALHALLRMCTMARHLAGRYVQVMAALALGRMTGYIYDHLTPM 956
DB 897 GPLMVLQAGITTRVPEVFRALHALLRMCTMARHLAGRYVQVMAALALGRMTGYIYDHLTPM 956
QY 961 SPMASGLRDLAAVEPILFSPMEKKVYVWGAETAAQSDILHGLPVASRLGRBVLIGPAD 1020
DB 961 SPMASGLRDLAAVEPILFSPMEKKVYVWGAETAAQSDILHGLPVASRLGRBVLIGPAD 1020
QY 957 RDMPRAGLRDLAAVEPILFSPMEKKVYVWGAETAAQSDILHGLPVASRLGRBVLIGPAD 1016
DB 957 RDMPRAGLRDLAAVEPILFSPMEKKVYVWGAETAAQSDILHGLPVASRLGRBVLIGPAD 1016
QY 1021 GYTSKMSLLAPITAAOQTRGLGRTIVSMTRGRDTEBAGEIQLVSTYQSLGTSIG 1080
DB 1021 GYTSKMSLLAPITAAOQTRGLGRTIVSMTRGRDTEBAGEIQLVSTYQSLGTSIG 1080
QY 1017 SLBGRRLRLAPITAAOQTRGLGRTIVSMTRGRDTEBAGEIQLVSTYQSLGTSIG 1076
DB 1017 SLBGRRLRLAPITAAOQTRGLGRTIVSMTRGRDTEBAGEIQLVSTYQSLGTSIG 1076
QY 1081 VLMTVHAGANKTLIGSRGPVTOMYSSAEGDVLGMPSPGTSLEBCTGADVLYVTNN 1140
DB 1081 VLMTVHAGANKTLIGSRGPVTOMYSSAEGDVLGMPSPGTSLEBCTGADVLYVTNN 1140
QY 1077 VCMTVHAGANKTLIGSRGPVTOMYSSAEGDVLGMPSPGTSLEBCTGADVLYVTNN 1136
DB 1077 VCMTVHAGANKTLIGSRGPVTOMYSSAEGDVLGMPSPGTSLEBCTGADVLYVTNN 1136
QY 1141 ADVIPARRRGRDYGALLSPRLSTLKGSGGCVLCRGAHVGFRACVSCRYASIDPT 1200
DB 1141 ADVIPARRRGRDYGALLSPRLSTLKGSGGCVLCRGAHVGFRACVSCRYASIDPT 1200
QY 1137 ADVIPARRRGRDYGALLSPRLSTLKGSGGCVLCRGAHVGFRACVSCRYASIDPT 1196
DB 1137 ADVIPARRRGRDYGALLSPRLSTLKGSGGCVLCRGAHVGFRACVSCRYASIDPT 1196
QY 1201 PVETLIDIVTRSPSPSNDSPPAVPTQVGYLHAAPGSGSKTKVPAYAAQGVKVLVLP 1260
DB 1201 PVETLIDIVTRSPSPSNDSPPAVPTQVGYLHAAPGSGSKTKVPAYAAQGVKVLVLP 1260
QY 1197 PVESHTTMRSPVFTDNSSPPAVPOSFOVAHLHAAPGSGSKTKVPAYAAQGVKVLVLP 1256
DB 1197 PVESHTTMRSPVFTDNSSPPAVPOSFOVAHLHAAPGSGSKTKVPAYAAQGVKVLVLP 1256
QY 1261 SVAATLFGAYLSKAGHINPNIRGTGRTVTTGAPITVSYGKFLADGGCAGAYDIIICD 1320
DB 1261 SVAATLFGAYLSKAGHINPNIRGTGRTVTTGAPITVSYGKFLADGGCAGAYDIIICD 1320
QY 1257 SVAATLFGAYLSKAGHINPNIRGTGRTVTTGAPITVSYGKFLADGGCAGAYDIIICD 1316
DB 1257 SVAATLFGAYLSKAGHINPNIRGTGRTVTTGAPITVSYGKFLADGGCAGAYDIIICD 1316
QY 1321 ECHAVDSTTILGIGYVLDQAEFTAGVRLTYLATATPPGSYTTHPNIEVALGOEGEIPY 1380
DB 1321 ECHAVDSTTILGIGYVLDQAEFTAGVRLTYLATATPPGSYTTHPNIEVALGOEGEIPY 1380
QY 1317 ECHAVDSTTILGIGYVLDQAEFTAGVRLTYLATATPPGSYTTHPNIEVALGOEGEIPY 1376
DB 1317 ECHAVDSTTILGIGYVLDQAEFTAGVRLTYLATATPPGSYTTHPNIEVALGOEGEIPY 1376
QY 1381 GRAIPLSYKGGRRHILFCHSKKKKDELAALRGMLNSVAYYRGLDVSVIPTQGDVVVVA 1440
DB 1381 GRAIPLSYKGGRRHILFCHSKKKKDELAALRGMLNSVAYYRGLDVSVIPTQGDVVVVA 1440
QY 1377 GRAIPLSYKGGRRHILFCHSKKKKDELAALRGMLNSVAYYRGLDVSVIPTQGDVVVVA 1436
DB 1377 GRAIPLSYKGGRRHILFCHSKKKKDELAALRGMLNSVAYYRGLDVSVIPTQGDVVVVA 1436
QY 1441 TDALMTGYTGDPSVIDCNVATQVVDPSLDPTFTTQIVPQDAVSRQSRGRTRGR 1500
DB 1441 TDALMTGYTGDPSVIDCNVATQVVDPSLDPTFTTQIVPQDAVSRQSRGRTRGR 1500
QY 1437 TDALMTGYTGDPSVIDCNVATQVVDPSLDPTFTTQIVPQDAVSRQSRGRTRGR 1496
DB 1437 TDALMTGYTGDPSVIDCNVATQVVDPSLDPTFTTQIVPQDAVSRQSRGRTRGR 1496
QY 1501 GIYRYVSTGERASGMPDVSIVCECYDAGAAWELTPSETTVRLRAYENTPGLPVCODHLE 1560
DB 1501 GIYRYVSTGERASGMPDVSIVCECYDAGAAWELTPSETTVRLRAYENTPGLPVCODHLE 1560
QY 1497 GIYRYVSTGERASGMPDVSIVCECYDAGAAWELTPSETTVRLRAYENTPGLPVCODHLE 1556
DB 1497 GIYRYVSTGERASGMPDVSIVCECYDAGAAWELTPSETTVRLRAYENTPGLPVCODHLE 1556
QY 1561 FWEAVFTGLTHIDAFHLSQTKOAGNPFVAVVQATVCARAPPSWQMMKCLRLRP 1620
DB 1561 FWEAVFTGLTHIDAFHLSQTKOAGNPFVAVVQATVCARAPPSWQMMKCLRLRP 1620
QY 1557 FWEAVFTGLTHIDAFHLSQTKOAGNPFVAVVQATVCARAPPSWQMMKCLRLRP 1616
DB 1557 FWEAVFTGLTHIDAFHLSQTKOAGNPFVAVVQATVCARAPPSWQMMKCLRLRP 1616
QY 1621 TLVGPPLLYRLSGVTNEVTLTHPYTKYIATCMQADLEVMSTWVLAGVLAVAAYCIA 1680
DB 1621 TLVGPPLLYRLSGVTNEVTLTHPYTKYIATCMQADLEVMSTWVLAGVLAVAAYCIA 1680
QY 1617 TLHGPPLLYRLSGVTNEVTLTHPYTKYIATCMQADLEVMSTWVLAGVLAVAAYCIA 1676
DB 1617 TLHGPPLLYRLSGVTNEVTLTHPYTKYIATCMQADLEVMSTWVLAGVLAVAAYCIA 1676
QY 1681 TGCVCILGRHINQAAVAPDKEVLYEAFDEMEBCASRAALIEGORAIEMLKSKIOGIL 1740
DB 1681 TGCVCILGRHINQAAVAPDKEVLYEAFDEMEBCASRAALIEGORAIEMLKSKIOGIL 1740
QY 1677 TGSVYIVGRHILISGRPAIVPRELLELYOEFDEMEBCASHLPYIEQWOLAEQKORALGIL 1736
DB 1677 TGSVYIVGRHILISGRPAIVPRELLELYOEFDEMEBCASHLPYIEQWOLAEQKORALGIL 1736

```

Qy	174	QOAKOKAODILOPTVOASMPKXECQEWAMHNMFIIGIOYLIGLSTLPNGNPAAWMAASAA	1800
Db	1737	QTATKOEAAPAVVESKRALETWAKMMNFISGIOYLIGLSTLPNGNPALASMAATAS	1796
Qy	1801	LTPSLSTSTTLILNLIGWMLASQIAPAPGATFVSVGLVGAANVSGIGKVLVDIILGYG	1860
Qy	1917	NRLTAFASRGHNVPTHYVTESDASORXTOLGSLITTLILRRHNMITECCPIPCGGSW	1980
Db	1857	AGVAGALVAFKVMGEMPESTEDLVNLLPALISPALVGVCAALIRRHVPGEGAAVQM	1916
Qy	1981	LRDYVWVCTIITDPEKMWLTSKLEPPKMGFLPFVSCQKGYKVMAGTGMITTRCPCGANIS	2040
Db	1977	LRDVWDMICTYLTDPEKMWLTSKLEPPKMGFLPFVSCQKGYKVMAGTGMITTRCPCGANIS	2036
Qy	2041	GNVFLGSMRITGPRTCMNIWGTFFPINCYTEGQCVKPAENFKVALIRVAASEAYEYQH	2100
Db	2037	GHWNGSMRIVGPKTCSWTHGTPEPINAVTTGCTPSPAPYNSALIRVAASEAYEYQTRV	2096
Qy	2101	GSYHYITGLTNDLKNVQCLSPSEPFESWVQVQIHRAPRTKPFPRDSEVSCVNSFVY	2160
Db	2097	GDFHYVYGMTTDNKKCCQVPAPEFSEVDVRLHRVAPACRPLRLREYVTFQVGLNDYLV	2156
Qy	2161	GSQJPCDPEPDTYLMGMLTDPSHITETAARLARAGSPESASASSASQLSAPSLATCT	2220
Db	2157	GSQJPCBEPEPVAVLTSMLTDPSHITETAARLARAGSPESASASSASQLSAPSLATCT	2216
Qy	2221	THGKAYVDVNVANLF---MGDVTRIESGSKYVVLDSLDPMVEBSQLEPSISEYML	2276
Db	2217	THHVSPPADLLEANULMRQEWGGINTRVESENKVVVLDSFPLRLAEEDEREVSAPAEILR	2276
Qy	2277	PKKFPFPAALPMARPDVNPRLVSEMKRPDQVAPVAGALPEPKTTPPRPRRRRTYGLS	2336
Db	2277	KSKFPFPAALPMARPDVNPRLVSEMKRPDQVAPVAGALPEPKTTPPRPRRRRTYGLS	2336
Qy	2337	EDSIGDALMOQALAIKSFQPPSPGDSGLSTGAGADSGQTP-PEBLA-----LSETGSIS	2390
Db	2337	ESSVSSALAEIATITGFC-----SSESAADSGATYALPQASDQDGKGDVESYS	2386
Qy	2391	SMPLLEGBELGDPLEBQEVQPPQGVAAFGSDSGSWSTCSSE-DSVVCSSMSYWT	2449
Db	2387	SMPLLEGBEPDPL-----SDGSWSTVSEASBDVVCSSMSYWT	2426
Qy	2450	GALITPSPSEBEKPIPNPLNSILRRHNMVYCTTTKASLSAPAKATYPRMAYLDSYDSV	2509
Db	2427	GALITTPCAABESKPIPNALNSILRRHNMVYATTSRSAGLRQKKTVPDRLOVLDHXYDV	2486
Qy	2510	LKQIKLAAKKTATALLTMEBAQCOTLTPPSAKSKYGFGEKVEYSLSGRAVNHIKSVMDLL	2569
Db	2487	LKEHKAKASTYKAKLISVEEBCKITPPHSAKSKGYGAKDVRANLSSKAVNHHSVMXDLL	2546
Qy	2570	EDSETPIPTITIMAKNEVFCVDPRTGKGGKKAALYVYDLYGVRCERKMLYDITQKLPQAVM	2629
Db	2547	EDVTPIPTITIMAKNEVFCVQPEKGGKRRPARLIYFPDLGVRCERKMLYDVSTLQVYM	2606
Qy	2630	GASYGFOYSPQORVEFLIKAAEKKDPMGSEYDTRCPDSTYTERDITRESITYACSLPE	2689
Db	2607	GSSYGFOYSPQORVEFLVNTWKSKKNPMGSEYDTRCPDSTYTERDITRESIYOCCLAP	2666
Qy	2690	EAHRAHSLTERLYVCGPMFNSKQOTCGYRCGRASGVYITSMGNITCYVVALAACRAAG	2749
Db	2667	EAHQAKSLTERLYVCGPULNSKQNGCYRCGRASGVYITSCGNITCYLVALASACRAK	2726
Qy	2750	IAATMLVCGDDLVIYSESQGTBEDERNLRAFTAMTRYSAAPGDPRPEYDELILITSQS	2809
Db	2727	LQDCMTLVNGDDLVIVYCESAGTQEDBAASLRVFTEMTRYSAAPGDPRPEYDELILITSQS	2786
Qy	2810	SNSVVALGPOGRRRYLITRDTTPPIARAAMETVRHSPVNSWLGNIIOYAPITIMARVLMAT	2869

Db	2787	SNVSVAHSDSGKRVYLLTRDPTPLAKAAWTAHTITVNSLGNIIIMYATLTAARIILMT	284
Qy	2870	HFFSILMAODTLDDINLFEMYGAVYSSPLDPAIIERLHGLDAFSLHTYTHELTRVAS	292
Db	2847	HFFSILMAOELEKALDCOIVGACYSIPDLPLQIIRLHGLSAFSLHYSPEIINRVAS	290
Qy	2930	ALMKLGAPLRLAMKSRRAAAYASLSLRGGAAVCGRTLFWMAVYTKLKLPLDEALLD	298
Db	2907	CLFKLGVPPLRLRWRRHRSVRRRLISQGGRAATGKTLFWMAVYTKLKLPLPAASRLD	296
Qy	2990	SSWFTVGAGGDIYHSVSRARPRLLFLGLLLPVGVLFLPLAR	3033
Db	2967	SGHFVAGYSGDIYHSLSRARPRMCLLLLSVGVLFLPLNR	3010
RESULT 15			
US-09-539-601-33			
; Sequence 33, Application US/09539601C			
; Patent No. 6630343			
; GENERAL INFORMATION:			
; APPLICANT: Barteneschlager, Ralf FW			
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System			
; FILE REFERENCE: all sequences			
; CURRENT APPLICATION NUMBER: US/09/539,601C			
; CURRENT FILING DATE: 2001-08-30			
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY			
; EARLIER FILING DATE: 1999-04-03			
; NUMBER OF SEQ ID NOS: 51			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 33			
; LENGTH: 3010			
; TYPE: PR1			
; ORGANISM: Hepatitis C virus			
US-09-539-601-33			
Query Match 74.3%; Score 11984.5; DB 4; Length 3010;			
Best Local Similarity 71.5%; Pzed. No. 0;			
Matches 2174; Conservative 353; Mismatches 477; Indels 35; Gaps 7			
Qy	1	MSTNPPROKTKRNTNRPRPDYKFPFGGQIVGVYLLPRGRPLRGVATRKTSERSQPRG	60
Db	1	MGTNPPROKTKRNTNRPRPDYKFPFGGQIVGVYLLPRGRPLRGVATRKTSERSQPRG	60
Qy	61	RROPIDKRRSTGKSGKGPYWPPLYNGEGLGMAWLLSPRGSRPSPGNDPRHRSRNVG	120
Db	61	RROPIDKRAQPRGRAMAQGYWPLVNGEGLGMAWLLSPRGSRPSPGNDPRHRSRNVG	120
Qy	121	KVIDTLTTCGADLMGTYIPVNGAPLPGVAKALAHGRVLVEGQVNPATNLPFGCSFIFLLA	180
Db	121	KVIDTLTTCGADLMGTYIPVNGAPLPGVAKALAHGRVLVEGQVNPATNLPFGCSFIFLLA	180
Qy	181	LWSCTTPPSAAEYKNVISGTGYVNTDCTNDSITWQOLAALVHPGVCPCEKVNASOCWI	240
Db	181	LWSCTTPPSAAEYKNVISGTGYVNTDCTNDSITWQOLAALVHPGVCPCEKVNASOCWI	240
Qy	241	PVSPNVAVORPGALTQGLRTIDWVMSATLCSALVYVDLCGVMLAQMFIVSPQHMF	300
Db	241	ALPFTLAARNAVPTTTIRRHVDLLVGAALCSAMYVDLCGVSFLVAQLFTPSPRRHE	300
Qy	301	VQDCNCSITPGTITGRMAAMDMNMWSTPAIMILAYMRVDEVYIIDSAGHGMVFGLA	360
Db	301	VQDCNCSITPGHVTGHRMAAMDMNMWSTPAIIVVOLLRIQAVVDVWAGAHGVLAGLA	360
Qy	361	YFPMOGAMKVVVILLILAAGVAPRHHTVGSAAOTGRLTSLPMPGMRKQICQVNTNGSW	420
Db	361	YFPMOGAMKAVLVMLFPAVGDDGTYTGTGMAKNTGIGTSLSPSSQKIQVNTNGSH	420
Qy	421	HIRRTALNCNDSLHTGTFLASLFTYTHSPFSSGCCPERMSACRSIEAFRGMGALDYEDNVN	480
Db	421	HIRRTALNCNDSLNTNGFLAALFYTHKFPSSGCCPERMASCSPIIDAFAGQWCPITY--NESH	478
Qy	481	PEDMRVCMHYPRQCGVSAKTVCGPVYCTPSPVAVGTTDRIGADTYTWGENETDVEL	540


```

Db 479 SSSDORPCYCHVA BRPGCIVPAQVCGPVYCFPTSPVAVGTTRDFGVPPTSMGSENFDDVL 538
Qy 541 LNSTRPPLGCSWFGCTMMNSGTYKTGAPPCRTADPNASTDLLCTDCCRPKPDPTLYLK 600
Db 539 LNNTRPPOQNMFGCTMMNSTGFTKTCGGPCPNIGIGNKT--LTCCTDCFRKRPBEATYTK 596
Qy 601 CGSGPMLPRCLIDYRPLMHVPCNTNYT PKIRMVGVGEHRLTPAANCNTRGDRCNLEB 660
Db 597 CGSGPMLPRCLVHYRPLMHVPCNTNYT FKVRMVGVGEHRLTPAANCNTRGDRCNLEB 656
Qy 661 RDRSOLSPLLHSTTEMAILPCSYSDLPALSTGLHLHONIVDOVMFGKSPALTKTYIV 720
Db 657 RDRSOLSPLLHSTTEMAILPCSYSDLPALSTGLHLHONIVDOVMFGKSPALTKTYIV 716
Qy 721 EBYVLLPFLLDARVCAIIMMLILLOAEALAEKVLILHAASACNGELVYVIFVVA 780
Db 717 EBYVLLPFLLDARVCAIIMMLILLOAEALAEKVLILHAASACNGELVYVIFVVA 776
Qy 781 YIKGRVPLATYSLNGLMSFSLILLALPOAAVAYDASVHGQI GALLNMTLFTLTPGYK 840
Db 777 YIKGRVPLATYSLNGLMSFSLILLALPOAAVAYDASVHGQI GALLNMTLFTLTPGYK 836
Qy 841 TILSRELMMWLCYLLTLGEAMVQEMAPMOVRGGRDGI IWAVALFYRPGVFDITKMLAVL 900
Db 837 LFLARILMWLOFYITRBAEHLQWIFPLAVNGRDAVILLTCAIHDELFTITKILLAL 896
Qy 901 GRAYILKALTRPVYFVRAHALLRMCOTMAHLAGRYVOMALLALGRMTGYIYDHLTPM 960
Db 897 GRLMVLQAGITTKVPYFVRAHALLRMCOTMAHLAGRYVOMALLALGRMTGYIYDHLTPM 956
Qy 961 SDMAASGLBDLVAVEPIIFSPMEKVIYWGAEATACGDIILGLPYSARLSREVLGPD 1020
Db 957 RDMAHAGLBDLVAVEPIIFSPMEKVIYWGAEATACGDIILGLPYSARLSREVLGPD 1016
Qy 1021 GYTSKMSMLABITLTAOQTRGLGTYIVSMGTGRDTEOAGEIOVLSTYOSFELGTSISG 1080
Db 1017 SLEGGMRLLAPITNVSQOTRGLGCIITSLTGDRDNQVGEVQVSTTQSLATCANG 1076
Qy 1081 VLMYTYHAGNKTLAGSRPVTOMTSSABGLVGMPSPPGTSLEPCTCGAIDLVLVTEN 1140
Db 1077 VCMYTYHAGNKTLAGSRPVTOMTSSABGLVGMPSPPGTSLEPCTCGAIDLVLVTEN 1136
Qy 1141 ADVITARRGDRGALLSPRLSTLKGSGGPRLCRGHANGVFRPAVCSRGVAKSIDRT 1200
Db 1137 ADVITARRGDRGALLSPRLSTLKGSGGPRLCRGHANGVFRPAVCSRGVAKSIDRT 1196
Qy 1201 PVEITDILVRSPTFSNSTRPPAVPOTYQVGYLHAPFGSGSKTKVPAAVYAGQIKVLVLP 1260
Db 1197 PVEITDILVRSPTFSNSTRPPAVPOTYQVGYLHAPFGSGSKTKVPAAVYAGQIKVLVLP 1256
Qy 1261 SVAAATLFGAYLSKAHGINPNI RTGVYVITGAPITYSTYKFLADGCGAGAYDIICD 1320
Db 1257 SVAAATLFGAYLSKAHGINPNI RTGVYVITGAPITYSTYKFLADGCGAGAYDIICD 1316
Qy 1321 BCHAVIDSTIILGIVLDOAEATAGVRLTYLATATPPGSAVTTPHPIEVALSTGSIPEY 1380
Db 1317 BCHAVIDSTIILGIVLDOAEATAGVRLTYLATATPPGSAVTTPHPIEVALSTGSIPEY 1376
Qy 1381 GRAIPLSYIKGRHLIFCHSKKKCDLAAALRGMGLNSVAYYRGDLVSIPTQGDVVVA 1440
Db 1377 GRAIPLSYIKGRHLIFCHSKKKCDLAAALRGMGLNSVAYYRGDLVSIPTQGDVVVA 1436
Qy 1441 TPLMTGYIGDSDVYDCNVAVTQVVDLSLPTITTOIIVODAVSRSQRRGRTGRGL 1500
Db 1437 TPLMTGYIGDSDVYDCNVAVTQVVDLSLPTITTOIIVODAVSRSQRRGRTGRGL 1496
Qy 1501 GYRYVSTGERASGMDSVVLCECYDAGAAWELPSETTVLRAVFNTPGLPYCODHLE 1560
Db 1497 GYRYVSTGERASGMDSVVLCECYDAGAAWELPSETTVLRAVFNTPGLPYCODHLE 1556
Qy 1561 FWEAVFTGLTHIDAFHLSQTKOSGENFAVLTAYQATVCARAKAPPEPSMDVMKCLTRLP 1620
Db 1557 FWEAVFTGLTHIDAFHLSQTKOSGENFAVLTAYQATVCARAKAPPEPSMDVMKCLTRLP 1616
Qy 1621 TLVGPRLYLXGVTNEVTLTHPVTXYATCMQADLEMTSTWYLAGVLAVAAYCLA 1680
Db 1617 TLHGPRLYLXGVTNEVTLTHPVTXYATCMQADLEMTSTWYLAGVLAVAAYCLA 1676
Qy 1681 TGCVCIIIGRLHINORVAVADKEVLYEAFDEMEBCASRAALJEBGORIAEMLSKIOGLL 1740
Db 1677 TGSVYIVGRIIISGRKAIIPDREVLVREDEMEBCASHLPYIEQSQMLBQFOKAIIGLL 1736
Qy 1741 QOASKOADIOPVQASWPKVEOPFAKHMNFIISGIOYLAGISTLPGNPAVSAWMAFSA 1800
Db 1737 QOATKOAEAAAVVESKMTIEAFMAKHMNFIISGIOYLAGISTLPGNPAVSAWMAFSA 1796
Qy 1801 LNSPLSTSTIILNIGWLASQIAPAGATGVVSGVGAAGVSGIGLKVTVLIDLAGV 1860
Db 1797 LNSPLSTSTIILNIGWLASQIAPAGATGVVSGVGAAGVSGIGLKVTVLIDLAGV 1856
Qy 1861 AGISGALVAFKIMSSEKPEMEDVNMILPGILSPGALVGVICAILRHHVGPGEAGVQM 1920
Db 1857 AGVAGALVAFKIMSSEKPEMEDVNMILPGILSPGALVGVICAILRHHVGPGEAGVQM 1916
Qy 1921 NRIIPASRGNHVAPTHVYTESDASQRYTOLGSLTISLRLHNMWITEDCPIPCGSGM 1980
Db 1917 NRIIPASRGNHVAPTHVYTESDASQRYTOLGSLTISLRLHNMWITEDCPIPCGSGM 1976
Qy 1981 LBDVMDVCTIITLDRKMLTSLCPKMRGLPVSQCKYKGYWAGTGITTCPCGANIS 2040
Db 1977 LBDVMDVCTIITLDRKMLTSLCPKMRGLPVSQCKYKGYWAGTGITTCPCGANIS 2036
Qy 2041 GNVRLGSMRITGPCKOMNI MOGTPPINCYTEOCQVCPKPNKVAIMRWASBYAEVTOH 2100
Db 2037 GNVRLGSMRITGPCKOMNI MOGTPPINCYTEOCQVCPKPNKVAIMRWASBYAEVTOH 2096
Qy 2101 GSXHYITGLTNDLVKPCQLPSPEEFSSWVQVQIHRFAPTPKPFREDEVSFCVGLNSFV 2160
Db 2097 GSXHYITGLTNDLVKPCQLPSPEEFSSWVQVQIHRFAPTPKPFREDEVSFCVGLNSFV 2156
Qy 2161 GSQLCDEPDPDVLMSMLTDSHITAEIARLARLARGSPSSASSASQLSAPSLATCT 2220
Db 2157 GSQLCDEPDPDVLMSMLTDSHITAEIARLARLARGSPSSASSASQLSAPSLATCT 2216
Qy 2221 THGKAYDVDMVANF---MGADVTRIESGSKVVLDSLDPMVERSPRLBSPSEYML 2276
Db 2217 THGKAYDVDMVANF---MGADVTRIESGSKVVLDSLDPMVERSPRLBSPSEYML 2276
Qy 2277 PKRFPFPALPAMARPDYNEPLVESWKRBDYQATVAGCALPPPKRTPTPPRRRRTVGLS 2336
Db 2277 PKRFPFPALPAMARPDYNEPLVESWKRBDYQATVAGCALPPPKRTPTPPRRRRTVGLS 2336
Qy 2337 EDSIGDALOQLAIKSPGPPPSG-DGSLSTGGAADSGSQTPPDELALSETGISIMPL 2395
Db 2337 EDSIGDALOQLAIKSPGPPPSG-DGSLSTGGAADSGSQTPPDELALSETGISIMPL 2391
Qy 2396 BEGLDPLDEPQVBPPOPGGVAAAPSDSGSMTSCSE--DSSVCCSMSTWMTGALT 2454
Db 2392 BEGLDPLDEPQVBPPOPGGVAAAPSDSGSMTSCSE--DSSVCCSMSTWMTGALT 2431
Qy 2455 PCSPEEKLPINPLNSILRYANKVYCTTYKASALRAKKTVPDRMQVLDSDYVSLYKDIK 2514
Db 2432 PCAAETKLPINALNSILRHHNLVYATTSRSANLRQKKVTDRLOVDLHDHRYDLKEMK 2491
Qy 2515 LAASKYTABLLTMEACQUTLPHSARSKYGFBAKAVRSLSGAANVHHSVMKDLLEDST 2574
Db 2492 LAASKYTABLLTMEACQUTLPHSARSKYGFBAKAVRSLSGAANVHHSVMKDLLEDST 2551
Qy 2575 PIPTIMAKNEVFCVDPPTKGGKARLIVYPLGVRVCEKMLYDITQGLPOAWGASYG 2634
Db 2553 PIPTIMAKNEVFCVDPPTKGGKARLIVYPLGVRVCEKMLYDITQGLPOAWGASYG 2611
Qy 2635 FOYSPAQVREFLKMAAKKQBPMSGYDTRCFDSVTYTERDITBESIYRACGLPEBAHTA 2694
Db 2612 FOYSPAQVREFLKMAAKKQBPMSGYDTRCFDSVTYTERDITBESIYRACGLPEBAHTA 2671

```

```

QY 2695 IHSILTERLYVGGPMFNSKQTCGYRCRASGVLTTSMGNTITCYVKALAAKAGIAPT 2754
Db 2672 IRSILTERLYTGGPLTNSKQNCGRRCRASGVLTTSCGNTLTCYLKAAAACRAKLODCT 2731
QY 2755 MLVCGDDLVIISQSGTEEDERNIRAFTEAMTRYSAAPPDPPREYDLELITSCSSNVSV 2814
Db 2732 MLVCGDDLVIICBSAGTQEDDEASIRAFTEAMTRYSAAPPDPPREYDLELITSCSSNVSV 2791
QY 2815 ALGPGRRRYLITRDPTTPIARAAWETVRHSVPVNSWLGNIIOYAPTIWARMVIMTHFESI 2874
Db 2792 AHDASGRKRYLITRDPTTPIARAAWETVRHSVPVNSWLGNIIMYAPTIWARMVIMTHFESI 2851
QY 2875 LMAQDTLDQNLNFMEMGAVYSVSPDLPAIIERLHGLDAPSLHTYTPHELTRVASALRKL 2934
Db 2852 LLAQEQLEKALDCQIYGACYSIBPLDLPQIIQRLHGLSAPSLHSYSPGEINRVASCLRKL 2911
QY 2935 GAPPLRAWKSRARAVRASLSIRGGRVAVCGRYLFNMAVTKTKLTPLEARLDDLSWFT 2994
Db 2912 GVPPLRYWRHRARSVRARLLSOGGRAATCGKYLEFNMAVTKTKLTPIPAASQDLDSWFT 2971
QY 2995 VGAGGDIYHSVRARPRLIFGLLLFVGVLFLPAR 3033
Db 2972 AGISGDIYHSLSRARPFWMCLLLSVGVIYLPNR 3010

```

Search completed: October 28, 2005, 16:07:10
 Job time : 83 secs